University of Bath

PHD

Theoretical modelling of organic reaction mechanisms in solution

Ruggiero, Guiseppe D.

Award date:
1999

Awarding institution:
University of Bath

Link to publication

Alternative formats
If you require this document in an alternative format, please contact:
openaccess@bath.ac.uk

Copyright of this thesis rests with the author. Access is subject to the above licence, if given. If no licence is specified above, original content in this thesis is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC-ND 4.0) Licence (https://creativecommons.org/licenses/by-nc-nd/4.0/). Any third-party copyright material present remains the property of its respective owner(s) and is licensed under its existing terms.

Take down policy
If you consider content within Bath's Research Portal to be in breach of UK law, please contact: openaccess@bath.ac.uk with the details. Your claim will be investigated and, where appropriate, the item will be removed from public view as soon as possible.

Download date: 05. May. 2023

Theoretical Modelling of Organic Reaction Mechanisms in Solution

by Giuseppe D. Ruggiero

Theoretical Modelling of Organic Reaction Mechanisms in Solution

Submitted by Guiseppe D. Ruggiero
for the degree of PhD
of the University of Bath
1999

COPYRIGHT

Attention is drawn to the fact that copyright of this thesis rests with its author. This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with its author and that no quotation from the thesis and no information derived from it may be published without the prior written consent of the author.

G. D. Ruggiero

UMI Number: U536700

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.

UMI U536700
Published by ProQuest LLC 2013. Copyright in the Dissertation held by the Author.
Microform Edition © ProQuest LLC.
All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

UNIVERSITY OF BATH
LIBRARY
30 - 7 FEB 2000
Ph.D.

Acknowledgements

I would firstly like to thank my supervisor, Professor Ian Williams who let me loose on the world of computational chemistry; without his guidance and lengthy interesting chemistry chats, none of the following work would have been possible. He has shown me that a Ph.D. is not solely about academic research.

Dr. Alex Turner and Dr. Vicent Moliner for their introduction and help with QM/MM modelling. (Now the student has become the Master!) The rest of the group; Stuart Firth-Clark, with whom I started and finished the life sentence that seemed a Ph.D. and Gail Rickard who would listen to my *excellent* singing and ranting and not complain Not forgetting Anders Holgner, aka the Swedish meatball lover.

Thanking Dr. Wayne Leonardo Shenton for his stimulating caffeine and alcohol diversions; long may it continue. Other people to mention who enriched my time at Bath; Mikey, Spunki, Taggart, Captain Pugwash, half-bro Moe, Wallop, Nigel, Iggy, Biggy Smalls, Tupac, Bacon, and last but definitely not least Lenny Draköt.

I would specially like to thank my Mum and Dad and the rest of my family who let my do what I wanted with my life. The biggest thanks of all goes to Claire and my son Alessandro, the people who give purpose and meaning to it all.

Summary

Theory and experiment are complementary tools in the search for fundamental understanding of chemical behaviour. Computational modelling has been applied to a variety of areas concerning organic reactivity, using a range of quantum chemical techniques from the simple AM1 semi-empirical Hamiltonian method to a more sophisticated level, such as the *ab initio* hybrid quantum-mechanical/molecular-mechanical method (QM/MM).

The semi-empirical continuum energy surfaces have been calculated for a series of increasingly alkylated protonated alcohols and for a series of substituted benzyl chlorides to investigate the point of mechanistic change for nucleophilic reactions. Saddle point searching and analysis of intrinsic reaction co-ordinates are used to model nucleophilic substitution reaction in solution with semi-empirical and *ab initio* based QM/MM techniques. The solvolysis of a series of alkyl chlorides, $H_2O + RCl \rightarrow ROH_2^+ + Cl^-$ (R = methyl, ethyl, *iso*-propyl, *t*-butyl and MeO), have been investigated using these techniques. The calculated kinetic isotope effects from these techniques were compared to experimental data in order to see if they can be used as reliable measures of transition-state structure and mechanism.

Theoretical modelling of transition-states can assist with the interpretation of experimental linear free energy relationships such as Hammett and Brønsted correlations. Theoretical modelling allows the calculation of all the steps in a reaction profile, whereas experimentally only the rate determining step can be seen. Theoretical semi-empirical continuum Hammett correlations have been calculated and compared to experimental correlations in order to determine their meaning. The transfer of the methoxycarbonyl group between isoquinoline and substituted pyridines in aqueous solution has been investigated using semi-empirical continuum methods and the results were compared with experiment.

Topological electron-density analysis of strained three-membered rings has been investigated using Bader 'atoms in molecules' analysis at the *ab initio* wavefunction level of theory.

These studies have shown the benefit of computational methods in studying complex reactions at the molecular level.

Table of contents

Chapter 9 Bimolecular Transition-states for Simple Alkyl Chlorides

Chapter 10 The Solvolysis of Methoxymethylchloride

Chapter 11 Transition-state Structural Refinement with GRACE and CHARMM: Modelling of Glycoside Hydrolysis in Aqueous Solution Using a Combined Quantum/Classical Method

Chapter One

Introduction

Chapter 1 - Introduction

1.1 Introduction

Computational chemical modelling attempts to improve understanding of known experimental results and predict the behaviour of unknown chemical systems. Many chemical reactions have no experimental probes, or such probes are complex and their interpretation controversial. Therefore, the search for organic reaction pathways for 'real' reactions using computational techniques can be extremely beneficial. A knowledge of a reaction mechanism can lead to rationalisation, correlation, and simplification of organic chemistry. The knowledge of the way known reactions proceed can enable us to predict the possible behaviour of related unknown reactions. If we can understand the energetics of a reaction, it may be possible to modify its course, i.e. to accelerate it, increase its selectivity or improve its yields. In drug design, it may be possible to build a mimic for the transition-state of a reaction normally catalysed by an enzyme. This 'transition-state analogue' binds to the receptor site of the enzyme which would normally interact with the actual transition-state for the reaction therefore disabling the enzyme. Once a transition-state for a catalysed reaction has been found and understood, it may be possible to test the effect that altering substituents has on the reaction.

1.2 Aims and objectives

The aim of this thesis is to show that modern computational techniques can be used to help understand and interpret fundamental chemical processes in solution. A mixture of *ab initio* and semi-empirical levels of theory is used along side continuum and hybrid quantum mechanical/molecular mechanical descriptions of the solution. This is achieved by completing the following objectives:

- Examining the effect of solvation upon the mechanism of a reaction.
- Interpreting secondary kinetic isotopes effects.
- Assessing structure-reactivity relationships.
- Evaluating transition-state models.

1.3 Overview of the thesis structure

1.3.1 Background

Chapter 2 deals with the theoretical models used in this thesis while Chapter 3 deals with reactivity probes.

1.3.2 Nucleophilic substitution

Nucleophilic aliphatic substitution was one of the first types of reaction whose mechanisms were subjected to detailed experimental study. During the 1930's much of the mechanistic background knowledge was assembled and interpreted by Ingold and Hughes. In Chapter 4, these pioneering mechanistic studies are revisited using modern computaional techniques. The influence of solvation upon the nucleophilic substitution on the degenerate reaction $H_2O + ROH_2^+ \rightarrow ROH_2^+ + H_2O$ (R = Me, Ethyl, ^{i}Pr and ^{t}Bu) is investigated using a continuum model to simulate the role of the solvent. At the beginning of each chapter a literature review of the subject matter is given.

1.3.3 Structure-reactivity relationships

Structure-reactivity relationships, such as the Hammett and Brønsted linear free energy relationships, have often been used as a means of correlating structure and reactivity, and as a mechanistic tool to probe the properties of the transition-state. The Hammett correlation is used to investigate the reactivity of benzyl substrates in Chapters 5 and 6, while in Chapter 7 the Brønsted correlation is used to investigate the transfer of the methoxycarbonyl group between isoquinoline and substituted pyridines.

1.3.4 Transition-state models

The transition-state structural changes can be rationalised at a descriptive level using two dimensional reaction co-ordinate diagrams and a simple model to explain this shifts (More O'Ferrall-Jencks). A valence bond configuration mixing model can also be used which is generated by a linear combination of valence bond electronic configurations (Pross-Shaik). In Chapter 8, nucleophilic substitution reactions at benzyl derivatives are investigated, and the results compared to theory.

1.3.5 Quantum mechanical and molecular mechanical modelling (QM/MM)

Historically, computational chemistry has treated quantum mechanics and molecular mechanics separately. Combined quantum/classical modelling is a computational method which allows calculations of chemical properties for very large systems. Solvent molecules are able to form hydrogen bonds to the reactants, transition-state or intermediates as a reaction progresses. The effect of solvation on a reaction can be to reduce the activation energy by preferentially stabilising the transition-state, or increase it by lowering the ground-state of the reactants. The role of specific solvent molecules is investigated using a quantum mechanical/molecular mechanical technique that involves transition-state structural refinement using the GRACE software package. QM/MM modelling is used to investigate the bimolecular reaction of a simple series of alkyl chlorides (Chapter 9), the solvolysis of methoxymethylchloride (Chapter 10), and glycoside hydrolysis in aqueous solution (Chapter 11). The calculated kinetic isotope effects are compared to experimental data in order to evaluate the computational models used.

1.3.6 Bader analysis

The structural properties of three membered rings are addressed using Bader's 'atoms in molecules' theory to compute their electronic topological characteristics. The three membered rings that are addressed in Chapter 12 are cyclopropane, cyclopropanone, oxirane, oxiranone and hydroxyoxiranone.

1.3.7 Conclusions and future work

Chapter 13 presents the conclusions of this thesis and discusses areas for possible future work.

Chapter Two

Concepts in Molecular Modelling

Chapter 2- Concepts in Molecular Modelling

2.1 Quantum mechanics

2.1.1 Ab initio

2.1.1.1 Introduction to molecular orbital calculations

Theoretical investigations throughout this thesis have been carried out using molecular orbital methods.[1] These quantum mechanical calculations provide approximate solutions to the non-relativistic time-independent Schrödinger equation:

$$H\Psi = E\Psi \qquad \text{Equation 2.1}$$

where H is the Hamiltonian operator that represents the sum of the kinetic and potential energy of the system, Ψ is the wavefunction, and E is the energy of the system. The Hamiltonian operator, H, includes terms for the potential and kinetic energy of both the nuclei and the electrons:

$$H = \sum_{a>b} \frac{Z_a Z_b}{r_{ab}} - \sum_{a,b} \frac{Z_a}{r_{ai}} + \sum_{i>j} \frac{1}{r_{ij}} + \sum_i \frac{1}{2}\nabla_i^2 - \sum_a \frac{1}{2m_a}\nabla_a^2 \qquad \text{Equation 2.2}$$

where a and b are nuclei, i and j are electrons, $\nabla^2 = \partial^2/\delta x^2 + \partial^2/\delta y^2 + \partial^2/\delta z^2$, $1/r_{ai}$ is the coulombic attraction between nucleus a and electron i; $1/r_{ij}$ is the repulsion between electron i and j.

The Hamiltonian can be simplified using the Born-Oppenheimer approximation whereby the motion of the electrons is decoupled from the motion of the nuclei. The total wavefunction for the system can then be written as:

$$\Psi_{total}(\text{nuclei, electrons}) = \Psi(\text{electrons})\Psi(\text{nuclei}) \qquad \text{Equation 2.3}$$

The total energy equals the sum of the nuclear energy and electronic energy:

$$E_{total} = E(\text{electrons}) + E(\text{nuclei}) \qquad \text{Equation 2.4}$$

When the Born-Oppenheimer approximation is used, the nuclei are said to be fixed while the electrons possess motion. For each arrangement of the nuclei the Schrödinger equation is solved for the electrons alone in the field of the nuclei. The total energy can be calculated by adding a constant nuclear repulsion term to the electronic energy, E_e. The equation can be solved to determine properties of a given system is given by the electronic Schrödinger equation:

$$H_e\Psi_e = E_e\Psi_e \qquad \text{Equation 2.5}$$

It is usual for the solutions of the wavefunction to be orthonormal so that there is an overall probability of 1 that the particle exists somewhere in any particular state, i.e.

$$\int \Psi_m^* \Psi_n d\tau = \delta_{mn} \qquad \text{Equation 2.6}$$

where $d\tau$ indicates that the integration is over all space, δ_{mn}=1 when m=n and 0 when m≠n.

The electronic wavefunction may be approximated as a product of the molecular orbitals, ϕ, which include a spin function, e.g. for a Li atom:

$$\Psi = \left| \phi_{is} \overline{\phi_{1s}} \phi_{2s} \right| \qquad \text{Equation 2.7}$$

where a function written without a bar is 'spin-up', whereas an over-lined function corresponds to 'spin-down'. The wavefunction must be anti-symmetric and the electrons indistinguishable (Pauli principle). Anti-symmetry is ensured by arranging the orbitals in a determinant known as a Slater determinant.

$$\Psi = \frac{1}{\sqrt{6}} \begin{vmatrix} \phi_{1s}(1) & \overline{\phi}_{1s}(1) & \phi_{2s}(1) \\ \phi_{1s}(2) & \overline{\phi}_{1s}(2) & \phi_{2s}(2) \\ \phi_{1s}(3) & \overline{\phi}_{1s}(3) & \phi_{2s}(3) \end{vmatrix} \qquad \text{Equation 2.8}$$

where $1/\sqrt{6}$ = normalisation constant. The average value of the energy can be calculated by multiplying Equation 2.5 by Ψ_e and integrating over all space and spin co-ordinates:

$$E_e = \frac{\int \Psi_e^* H_e^* \Psi_e^* d\tau}{\int \Psi_e^* \Psi_e^* d\tau} = \frac{<\Psi^* | H | \Psi>}{<\Psi^* \Psi>} \qquad \text{Equation 2.9}$$

2.1.1.2 The Hartree-Fock approximation

In accordance with the variational principle the 'best' wavefunction is obtained when the energy is at its minimum. At the minimum, the first derivative of the energy $\delta E=0$. The Hartree-Fock equations are obtained by imposing this condition on the energy expression, with the added constraint that the molecular orbitals remain orthonormal. The Hartree-Fock approximation treats the electrons as independent particles which move in an averaged field created by the nuclei and other electrons. A many electron system can then be treated as a series of single electron problems. The Hartree-Fock equation is given as:

$$f_i \phi = \varepsilon \phi \qquad \text{Equation 2.10}$$

where f_I is the Fock operator (an effective one-electron Hamiltonian for the electron in a polyelectronic system) having the form:

$$f_i = H_{(i)}^{core} + \sum_{ij}^{N}(J_{ij} - K_{ij}) \qquad \text{Equation 2.11}$$

where J_{ij} is the Coulomb operator (corresponding to the average potential due to an electron in ϕ_j, and K_{ij} is the exchange operator (relating to the exchange repulsion between electrons with the same spin). $H_{(i)}^{core} = -\frac{1}{2}\nabla_i^2 - \sum_{a=1}\frac{Z_a}{r_{ia}}$ (corresponding to the motion of a single electron moving in the field of bare nuclei) where ∇^2 is the position vector through space, Z_a represents a single electron with a single nucleus and r being the eigenvalue equal to a.

For a closed shell system, the Fock operator has the form:

$$f_i = H_{(i)}^{core} + \sum_{ij}^{N/2}(2J_i - K_i) \qquad \text{Equation 2.12}$$

A direct solution of the Hartree-Fock equations is not practical for molecules. They are solved iteratively using the self-consistent field (SCF) approach. In this approach a trial set of solutions for the molecular spin orbitals are obtained for the Hartree-Fock equations. These are used to calculate the exchange and Coulomb operators. The Hartree-Fock equations are then solved for a new set of molecular spin orbitals, which are used in the next iteration. The procedure is repeated until the results for all the electrons no longer change, when they are said to be self-consistent.

2.1.1.3 The Roothaan-Hall equations

Molecular spin orbitals can be written as a linear combination of single electron orbitals:

$$\Psi_i = \sum_{v=1}^{K} c_{vi}\chi_v \qquad \text{Equation 2.13}$$

The one-electron orbitals, χ_v, are basis functions and often correspond to atomic orbitals. The problem of calculating the lowest energy wavefunction can be reduced to determining the set of expansion coefficients c which derive that wavefunction. The best set of coefficients is that for which the energy is a minimum, i.e.

$$\frac{\partial E}{\partial c_{vi}} = 0 \qquad \text{Equation 2.14}$$

for all coefficients c when substituted into Equation 2.9. Roothaan and Hall[2] recast the Hartree-Fock equations in matrix form for closed-shell systems. The Fock matrix can be written as:

$$\sum_j F_{ij}c_{ij} = \varepsilon \sum_j S_{ij}c_{ij} \qquad \text{Equation 2.15}$$

where $F_{ij} = \langle\chi_k|F_i|\chi_j\rangle$, and $S_{kj} = \langle\chi_k|\chi_j\rangle$ = overlap integral. This can be written as the matrix equation:

$$FC = SCE \qquad \text{Equation 2.16}$$

where C is a square matrix of the coefficients c_{vi}, E is a diagonal matrix whose elements are the orbital energies, and S is the overlap matrix.

Standard matrix eigenvalue methods can only be used to solve the Roothaan-Hall equations if the overlap matrix, S, is equal to the unit matrix, I (where all the diagonal elements are equal to 1 and all off-diagonal elements are zero). As the functions χ are usually normalised but not necessarily orthogonal there will be non-zero off diagonal elements of the overlap matrix. The basis functions are therefore transformed to form an orthonormal set. This is achieved using a transformation matrix, X, such that:

$$X = IS^{-1/2} \qquad \text{Equation 2.17}$$

The transformation matrix must also act on the Fock and coefficient matrices to leave the Roothaan-Hall equations as eigenvalue problems:

$$F'C' = C'E \qquad \text{Equation 2.18}$$

where $F' = S^{-1/2}FS^{-1/2}$ and $C' = S^{-1/2}C$. The Roothaan-Hall equations can then be solved using standard iterative methods.

2.1.1.4 Basis Sets

Consideration must be given to the type of basis set used for *ab initio* quantum mechanical calculations on molecular systems. The basis sets most commonly used are composed of atomic functions. There are two main criteria used for choosing basis sets. The basis set must be capable of describing the actual wavefunction well enough to give chemically useful results and they must also allow the evaluation of the F_{kj} and S_{kj} integrals accurately and cheaply. The exact expansion of Equation 2.13 would require an infinite basis set, which is not possible. Smaller sets must therefore be chosen whose results give a compromise between the two criteria stated.

Two major types of function are used in *ab initio* molecular orbital calculations. These are Slater type orbitals (STOs) and gaussian functions. The solution of the Schrödinger equation for hydrogen-like atoms suggests the use of atomic orbitals of the form:

$$\psi = R(r)Y(\theta,\phi) \qquad \text{Equation 2.19}$$

where Y is the spherical harmonic and R is the radial function. STOs are derived from hydrogen atomic orbitals and have the form:

$$R(r) = Nr^{n-1}e^{-\xi r} \qquad \text{Equation 2.20}$$

However, the integrals related to these functions are difficult to evaluate especially when the atomic orbitals are centred on different nuclei. Hence STOs have been replaced by gaussian functions which lead to simpler integral equations. Gaussian functions have the general form:

$$R(r) = r^{n-1}\exp(-\alpha r^2) \qquad \text{Equation 2.21}$$

where α determines the radial extend or 'spread' of the gaussian function. The product of two gaussian functions is another gaussian function that is located along the line joining the centres of the two gaussian centres. The major differences between the two types of function are that gaussians have no cusp at the nucleus, and they die away more quickly at larger values of *r*. These deficiencies are important since they mean that the gaussian function does not resemble very closely the form of real atomic orbital wavefunctions. To minimise these problems, more than one gaussian can be used to describe a STO.

A STO can be represented by a series of gaussians, e.g.:

$$R(r)_{is} = \left| c_1 e^{-\alpha\zeta} + c_2 e^{-\alpha\zeta^2} + c_3 e^{-\alpha\zeta^3} \right| \qquad \text{Equation 2.22}$$

where c_1, c_2, and c_3 are contraction coefficients, and α_1, α_2, and α_3 are exponents. In this case 3 gaussian functions represent 1 STO. Both the contraction coefficients and the exponents are optimised in the atom to approximate a Slater 1s orbital. Alternatively, you can use a linear combination of gaussians to give double-zeta or split valence shell basis sets. These include the 3-21G and 6-31G* basis sets. The 6-31G* set also includes *d*-polarisation on second row elements.

Different basis sets produce different wavefunctions and energies. As the mathematical flexibility of the basis set is extended so the SCF energy becomes lower. The energy approaches a limiting value as the basis set approaches mathematical completeness. The limiting energy is the lowest that can be achieved for a single determinant wavefunction. This is termed the Hartree-Fock energy. The Hartree-Fock energy is not as low as the true energy of the system. To improve on the Hartree-Fock energy post-SCF methods can be used.

2.1.1.5 Electron correlation

However good the Hartree-Fock wavefunction is, it is not the 'true' wavefunction. Hartree-Fock relies on averages and so it does not take into account the instantaneous electrostatic interactions between electrons nor does it take into account the quantum mechanical effects on electron distribution because the effect of n-1 electrons on an electron of interest is treated in an average way. The Hartree-Fock method yields a finite set of spin-orbitals when a finite basis set expansion is used. In general, a basis with M members results in 2M different spin-orbitals. By ordering spin-orbits energetically and taking the n lowest in energy, the Hartree-Fock wavefunction Φ_0 is formed. However, there remains $2M$ - n virtual orbitals. Many Slater determinants can be formed from the 2M spin-orbitals and Φ_0 is just one of them. By using Φ_0 as a reference, it is possible to classify all other determinants according to how many electrons have been promoted from occupied orbitals to virtual orbitals. A single excited determinant corresponds to one for which a single electron in occupied spin-orbital Φ_3 has been promoted to a virtual Φ_4, Figure 2.2. A doubly excited determinant corresponds to one in which two electrons have been promoted, one from to Φ_3 to Φ_4, and one from to Φ_3 to Φ_5. In a similar manner, we can form other multiply excited determinants. Each of the determinants, or a linear combination of a small number of them is called a configuration state function. The excited configuration state functions can be taken to approximate excited-state wavefunctions or as a linear combination to improve the representation of the ground-state (or any excited-state) wavefunction.

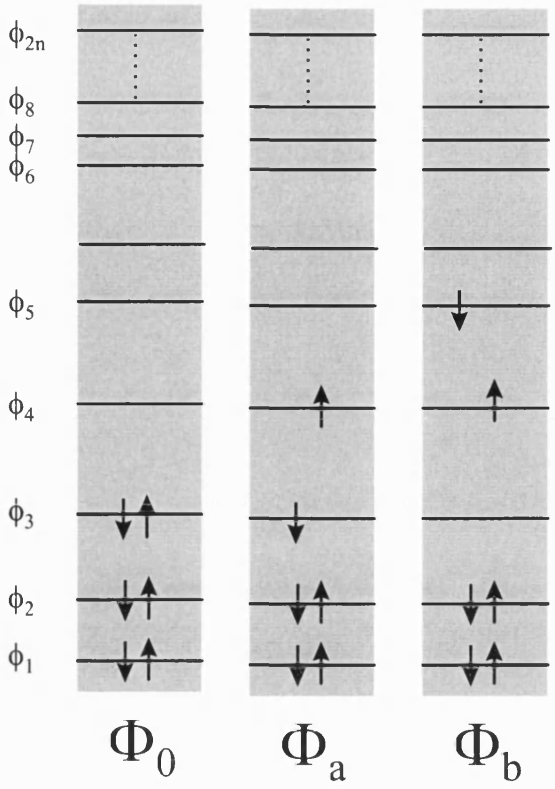


Figure 2.1: The notion for excited determinants.

2.1.1.5.1 Configuration interaction

The exact ground-state and excited-state wavefunctions can be expressed as a linear combination of all possible *n*-electron Slater determinants arising from a complete set of spin-orbitals. Therefore it is possible to write the exact electronic wavefunction ψ for any state of the system in the form:

$$\psi = C_0\phi_0 + \sum_{\phi_3\phi_4} C_1\Phi_1 + \sum_{\phi_3\phi_4} C_2\Phi_2 + \sum_{\phi_3\phi_4} C_3\Phi_3 + \ldots \qquad \text{Equation 2.23}$$

where the *C*s are expansion coefficients and the limits in the summations ensure that we sum over all unique pairs of spin-orbitals in doubly excited determinants, triply excited determinants, etc. An *ab initio* method in which the wavefunction is expressed as a linear combination of determinants is called configuration interaction. The difference between this exact energy and the Hartree-Fock energy is called the correlation energy. It is not possible to handle an infinite basis set of *n*-electron determinants with each determinant constructed from an infinite set of spin-orbitals. Unfortunately, even for a small number of electrons, the total of determinants can be extremely large therefore in practice, Equation 2.23 will always be truncated.

2.1.1.5.2 Møller-Plesset many body perturbation theory

Configuration interaction calculations provide a systematic approach for going beyond the Hartree-Fock level, by including determinants that are successively singly excited, doubly excited, triply excited, etc., from reference configuration. Møller-Plesset[3] perturbation theory provides an alternative systematic approach to finding the correlation energy. The number of problems which can be solved exactly by quantum mechanics methods is not very large. If our problem is to solve:

$$H\Psi_i = E_i\Psi_i \qquad \text{Equation 2.24}$$

This does not appear to have an easy solution, whilst the following simpler related problem can be solved exactly:

$$H^0\Psi_i^0 = E_i^0\Psi_i^0 \qquad \text{Equation 2.25}$$

For example, H might be the Hamiltonian for a He atom, and the zero order problem with the superscript 0 might refer to two superimposed H atoms whose electrons do not interact. The idea is to write the Hamiltonian H as:

$$H = H^0 + \lambda H^1 \qquad \text{Equation 2.26}$$

where H^0 refers to the zero order perturbation, whose solution we know, and H^1 is the perturbation. Normally, the arbitrary part-timer λ is included so as to keep track of the orders of magnitude in the derivation. Perturbation theory aims to write solutions for:

$$H\Psi_i = E_i\Psi_i \qquad \text{Equation 2.27}$$

in terms of the zero order problem, thus

$$\Psi_i = \Psi_i^0 + \lambda\Psi_i^1 + \lambda^2\Psi_i^2 + \ldots \qquad \text{Equation 2.28}$$

$$E_i = E_i^0 + \lambda E_i^1 + \lambda^2 E_i^2 + \ldots \qquad \text{Equation 2.29}$$

We call these expansions perturbation expansions of the wavefunction and the energy, and the result correct to second order is:

$$E_i = E_i^0 + \lambda H_{ii}^1 + \lambda^2 \sum_{m \neq i} \frac{H_{im}^0 H_{mi}^0}{E_i^0 - E_m^0} + \ldots \qquad \text{Equation 2.30}$$

with a corresponding expression for ψ_i. In this expression,

$$H_{mi}^1 = \int \Psi_m^0 H^1 \Psi_i^0 d\tau \qquad \text{Equation 2.31}$$

the idea is to take the zero order problem as the Hartree-Fock model and the perturbation as the difference between the true Hamiltonian and the Hartree-Fock Hamiltonian.

2.1.2 Semi-empirical

The use of *ab initio* techniques is prohibitively computationally expensive for larger systems. The majority of the time taken to calculate a molecule using *ab initio* methods is spent computing and manipulating integrals. In order to allow large systems to be treated, semi-empirical approaches may be used in which these integrals are generally approximated or neglected thereby speeding up the calculation. One of the main approximations used to simplify the calculations is to neglect the explicit representation of the core electrons. The effect of these core electrons is included in the representation of the nuclear core. The valence electrons are calculated explicitly as they are the electrons that are most associated with chemical interactions. The semi-empirical approaches generally use basis sets that consist of Slater-type orbitals, which are orthogonal; this allows more approximations.

Most semi-empirical techniques involve a simplification of the matrices used to calculate orbital overlap. These techniques include (ZDO) zero-differential overlap, (CNDO) complete neglect of differential overlap[4], (INDO) intermediate neglect of differential overlap for one-centre integrals[5] and (NDDO) neglect of diatomic differential overlap[6]. These approaches are rarely used today due to the rapid improvement in computer power since they were introduced and the extremes of the approximations used. The earlier of these is the modified MINDO/3[7] method which, although similar in theory to INDO, includes more parameterisation from experimental data. The MINDO/3 approach had some significant limitations such as errors in the calculation of the bond angles and the heats of formation of unsaturated molecules become too positive. Another approach was then derived called the modified neglect of diatomic overlap (MNDO) method based on the NDDO method. This approach had significant advantages over the MINDO/3 approach as the use of monatomic parameters allowed the inclusion of different core-core repulsion terms. The range of elements was also increased especially when the approach was modified further to include *d*-orbitals.

The Austin Model 1 (AM1)[8] was developed by the Dewar group to eliminate the problems with MNDO which would arise from a tendency to overestimate the repulsion between atoms separated by distances approximately equal to their van der Waals radii. The strategy to fix this problem was to modify the core-core term using gaussian functions. Both attractive and repulsive gaussians were used. The attractive gaussians

were designed to overcome the repulsion directly and were centred in the region where the repulsions were too large. The repulsive gaussian functions were centred at smaller inter-nuclear separations. The AM1 parameters were derived from empirical results and chemical intuition.

2.2 Molecular mechanics (MM)

Due to the computational expense, it is impossible to study 'large' molecules using rigorous quantum mechanical methods. The definition of 'large' changes as computer power increases. One of the goals of molecular modelling is to simulate the three-dimensional properties and structure of the molecule being investigated. Molecular mechanics (MM) techniques use various approximations to simplify the calculations involved in this process. The algorithms used produce a mathematical model of the properties of a system. In MM these algorithms are based on the laws of classical physics. MM works by using an empirical fit to calculate, for instance, the potential energy surface of the molecule, thus creating a forcefield. The two important features of a forcefield are its simplicity and its transferability throughout a variety of systems. The accuracy of the forcefield produced is dependent on the quality of the potential functions and parameters used. There are several disadvantages with this model compared with a quantum mechanical model, the main one being that it does not use an explicit representation of the electrons and therefore can not be used to describe bond making and breaking.

In the MM representation, the molecules are represented by a collection of balls joined by 'springs'. The strength of these springs differs depending on the various bond strengths. The motion is determined by the laws of classical physics and is represented by a combination of the intra-molecular and inter-molecular forces acting on the system. When characterising an atom using a forcefield, extra data is required to describe the atom types in the molecule. This information is usually related to the hybridisation state and the environment in which the atom is located. The forcefield is calculated from the functions representing the bonded (through bond) and non-bonding (through space) interactions. The total potential energy, E_{total}, consists of the following terms:

$$E_{total} = E_{bond\ length} + E_{bond\ angle} + E_{torsion} + E_{non\text{-}bonded} \qquad \text{Equation 2.32}$$

where the energy is considered as consisting of the bond length stretching energy, $E_{bond\ length}$, the bond angle bending energy, $E_{bond\ angle}$, and the torsion angle energy, $E_{torsion}$, along with non-bonded interactions, $E_{non\text{-}bonded}$.

One general four-term functional form of the algorithm used in a MM force field is as follows:

$$E_{total} = \sum_{bonds} \frac{k_l}{2}(l_i - l_0)^2 + \sum_{angles} \frac{k_i}{2}(\theta_i - \theta_0)^2 + \sum_{torrions} \frac{V_n}{2}(1+\cos(n\omega - \gamma))^2 \qquad \text{Equation 2.33}$$

$$+\sum_{i=1}^{N}\sum_{j=i+1}^{N}[4\varepsilon_{ij}[(\frac{r_0}{r_{ij}})^{12} - (\frac{r_0}{r_{ij}})^{6}] + \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}}]$$

where N is the number of particles, n is the periodicity, l_i is the bond length, l_0 is the equilibrium bond length, θ_i is the bond angle, θ_0 is the equilibrium bond angle and ω is the torsion angle. The non-bonded interactions are described in terms the charges of i and j, q_i and q_j, and the separation between the charges, r_{ij}, the rest of the terms are constants.

The bond length stretching energy can be treated by a Morse function instead of a simple harmonic oscillator. This makes more physical sense as the bond stretching energy is only harmonic near equilibrium and deviates greatly from this as the bond length increases. The Morse function has the form:

$$E_{bond\ length} = D_e(1 - e^{-\omega\sqrt{\frac{\mu}{2D_e}}\, l_i - l_0}) \qquad \text{Equation 2.34}$$

where D_e is the dissociation energy, μ is the reduced mass, ω is the bond vibration frequency, l_i is the bond length and l_0 is the parameterised bond length. The bond vibration frequency, ω, is related to the bond stretching constant, k_l, by $\omega = \sqrt{(\frac{k_l}{\mu})}$ where μ is the reduced mass of the diatomic molecule A–B, given by $1/\mu = 1/m_A + 1/m_B$. Close to equilibrium, l_i-l_0 is small, so the exponential function can be expanded and truncated to yield Hooke's law:

$$E_{bond\ length} = \sum k_l (l_i - l_0)^2 \qquad \text{Equation 2.35}$$

The bond angle bending energy is usually approximated for ease of computation using a simple harmonic oscillator as follows:

$$E_{torsion}=\sum k_{\theta}(\theta_i-\theta_0)^2 \qquad \text{Equation 2.36}$$

where k_θ is a force constant and θ_0 is the equilibrium bond angle. The existence of barriers to rotation around chemical bonds is represented by a torsional term, $E_{torsion}$, that is a function of the torsional angle, ω:

$$E_{\omega}=\sum V_n\,(1+\mathrm{s}\cos n\omega)^2 \qquad \text{Equation 2.37}$$

where V_n is the rotational barrier height, n is the periodic of rotation and s = 1 for a staggered minimum and -1 for an eclipsed minimum.

The non-bonding interactions are distance dependent and are calculated over all atoms with a 1,4 or greater separation. These forces are considered as having van der Waals, E_{vdw}, and electrostatic components, E_{el}:

$$E_{non\text{-}bonded} = E_{vdw} + E_{el} \qquad \text{Equation 2.38}$$

The electrostatic component does not explicitly represent the effect of polarisation on the system although these effects may be included implicitly through the parameterisation used. The van der Waals interaction, E_{vdw}, consists of attractive long-range forces and repulsive short-range forces. The attractive dispersive forces are referred to as London forces which are caused by the attractive dipole-dipole interactions. These transient dipoles are caused by the fluctuations in electron positions within the electron cloud. The dipole induces a dipole in adjacent atoms, producing an attractive inductive effect. The induced dipoles are vector quantities that do not average to zero as the orientation of the induced dipole is always in the same direction as the permanent dipole. These interactions are given by the London formula for two atoms *i* and *j*:

$$E=\sum -\frac{3\alpha_i\alpha_j I_i I_j}{2\,I_i+I_j\;r^6} \qquad \text{Equation 2.39}$$

where r is the distance between nuclei i and j, I_i and I_j are there ionisation energies and α_i and α_j are there volume polarisabilities. The strength of this effect is proportional to both the volume polarisabilities and the ionisation energies of the atoms involved. Crucially, the force is related to the distance, r, by a relationship proportional to r^{-6} therefore the force becomes insignificant after a few Ångstroms.

The opposing repulsion mainly occurs due to the electron-electron repulsion in the molecule when the distances between the atoms are near their contact radii. The exchange repulsion force is related to the inability of electrons of the same spin to occupy the same region of space. The exchange effect increases the repulsion between pairs of electrons as they are forbidden to occupy the same space, the inter-nuclear region. This causes the inter-nuclear region to be electron deficient thus reducing the shielding between the two nuclei, which in turn repel each other. This term is usually taken, as being proportional to r^{12} as this is inexpensive to compute. The combination of the attractive and repulsive terms is usually written in the form of the 12,6 Lennard-Jones potential. The 12,6 Lennard-Jones potential between two atoms is:

$$E_{vdw} = \sum 4\varepsilon\left(\frac{r_m}{r}\right)^{12} - 2\left(\frac{r_m}{r}\right)^{6} \qquad \text{Equation 2.40}$$

where r_m is the minimum energy interaction distance, r is the distance between the atoms and ε is related to the potential energy well depth. This can be written in terms of the collision diameter, σ. This is related to the minimum energy distance as $r_m = 2^{1/6}\sigma$, producing the following equation:

$$E_{vdw} = \sum 4\varepsilon\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6} \qquad \text{Equation 2.41}$$

The electrostatic distribution is usually represented as partial atomic charges distributed throughout the molecule. The electrostatic interactions between a pair of point charges can be derived from Coulomb's law; The interaction between two molecules can be written as:

$$E_{el} = \sum \frac{q_i q_j}{4\pi\varepsilon_o \varepsilon_r r_{ij}^2} \qquad \text{Equation 2.42}$$

where q_i and q_j are the partial charges on atom centres i and j, r_{ij} is the distance between the charges i and j, ε_o is the permittivity of free space and ε_r is the permittivity of the molecule concerned, which is dependent on the solvent used. These forces are at the core of most MM forcefields but other forces are sometimes included. The largest of these are hydrogen bonding and 1,3 interactions called cross terms, which account for the deformation of bonds.

2.3 Potential energy surfaces

Reaction profiles illustrate the way in which the energy of a system changes as a function of the reaction co-ordinate. The reaction co-ordinate is a general term used to represent the nuclear reorganisation that takes place when reactants are converted to products. The energy term used can refer to any one of the different thermodynamic measurements: enthalpy, H, free energy, G, or internal energy, E. Examples of such profiles are illustrated in reactions where reactants, R, are converted to products, P, in a single step, Figure 2.2(a) and for a multi-step reaction, Figure 2.2(b), which passes through a high energy intermediate, I. The conversion from reactants to products leads the system through a high energy structure that is a point of maximum energy along the reaction co-ordinate. This is termed the transition-state, TS.

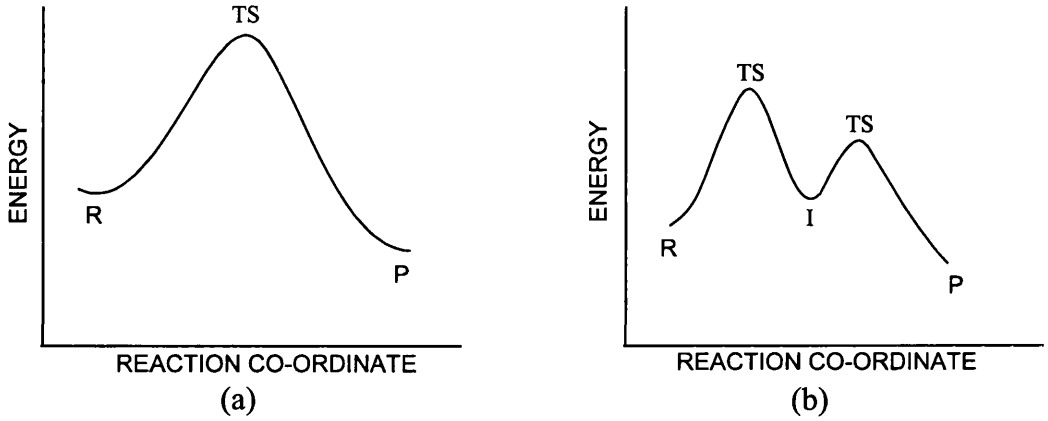


Figure 2.2: Energy profiles for (a) a single-step reaction proceeding from the reactants, R, through the transition state, TS, and down to the products, P. (b) a two-step reaction that proceeds through a high energy intermediate species I.

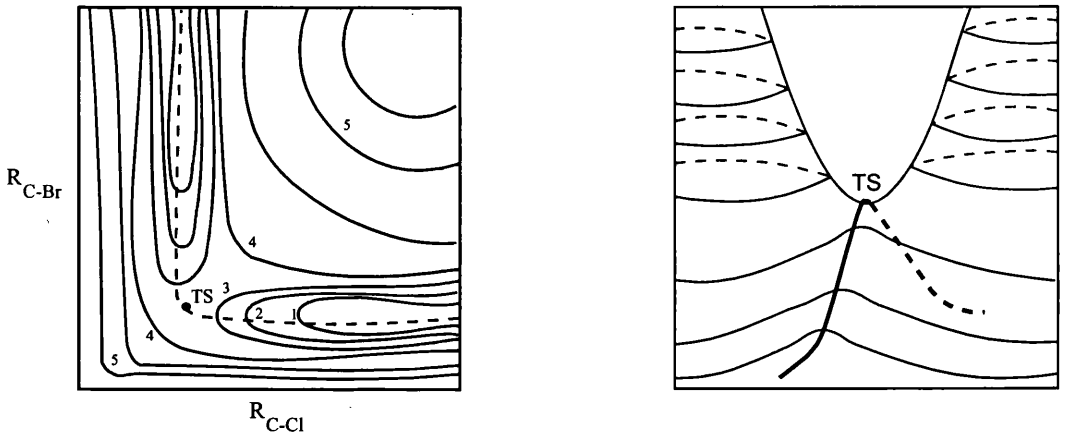
An alternative method for depicting reaction pathways is through the use of potential energy surface diagrams. These have the advantage of giving added geometric information as compared to the reaction profiles. A system of N atoms possesses $3N$ Cartesian degrees of freedom. Six of these degrees of freedom specify the absolute position and orientation of the system as a whole in Cartesian space: three define the position of the centre of mass of the system relative to the origin, while three (for a non-linear molecule) define the principle axes of the inertia of the system relative to the reference frame. The remaining $3N$ - 6 degrees of freedom define the possible internal vibrations of the system, i.e. the relative position of the atoms with respect to one another. The Born-Oppenheimer approximation, which states that the motion of electrons in a molecule is uncoupled from the motion of nuclei, means that any

molecular vibration can be defined in terms of these $3N$ - 6 internal degrees of freedom. (For a linear system there are $3N$ - 5 vibrational degrees of freedom since it possesses just two rotational degrees of freedom; the additional vibration takes the place of the missing rotation). Since in a chemical reaction we are interested in the way atoms change their position with respect to one another, chemical reactivity involves $3N$ - 6 degrees of freedom out of the total $3N$.

$$Br^- + CH_3Cl \rightarrow Cl^- + CH_3Br \qquad \text{Equation 2.43}$$

For example, consider a simple S_N2 reaction, Equation 2.43; this system consists of six atoms and therefore possesses twelve degrees of internal freedom (3 x 6 – 6 = 12). Consequently, the energy of the system is a function of twelve geometrical parameters, and a full graphical representation of this function requires thirteen dimensions and is therefore not readily visualisable. However, it is possible to construct a potential energy surface in a reduced space by considering the effect on the energy of only two degrees of freedom. The potential energy surface (PES) for reaction in Equation 2.43 can be seen Figure 2.3(a). The energy of the system is plotted as a function of two specific geometric parameters. These two parameters are in the plane of the paper with the third co-ordinate, the energy axis, perpendicular to the plane of the paper and the energy of the system indicated by the contour lines. Any line on the surface joins together points of equal energy, and the spacing between the contour lines signifies the gradient of the surface.

In principle, in drawing the PES any two of the twelve geometric variables may be chosen. However, since the two most important parameters are the two relating to the C–Cl and C–Br bond distances, it these two parameters that are chosen while the other parameters are relaxed. A 3D representation of the transition state region is illustrated in Figure 2.3(b). A characteristic of a saddle point is that it constitutes a maximum along the reaction co-ordinate and a local minimum for any geometric change perpendicular to the reaction co-ordinate. The saddle point corresponds to a transition-state on the PES. There are restoring forces opposing distortion in $3N$-7 vibrational modes perpendicular to the reaction path, but no restoring force opposing distortion parallel to the reaction path.



(a) (b)

Figure 2.3: (a) Schematic potential energy surface diagram for the S_N2 reaction, $Br^- + CH_3Cl \rightarrow Cl^- + CH_3Br$. The axes represent C–Cl and C–Br bond stretching and the third energy dimension is represented by contour lines. The reaction co-ordinate is represented by the dashed line. High energy contours are represent by 5 while low energy contours are represent by 1. (b) 3D representation of the saddle point region of the energy surface. The saddle point is the location of the transition-state, TS.

2.4 Solvation

2.4.1 Quantum mechanics/molecular mechanics (QM/MM)

Hybrid quantum mechanics/molecular mechanical modelling aims to combine the strengths of both the quantum mechanical and molecular mechanical methods. The QM/MM method possess a quantum mechanical core which is surrounded by a standard molecular mechanics potential and several interaction terms join the two. The atoms that are most influenced by electronic redistribution, in the property under investigation, are included in the quantum mechanical core with the bulk of the atoms placed in molecular mechanical surroundings. A more detailed description of hybrid quantum mechanics/molecular mechanics methodology is presented in Chapter 9.

2.4.2 Continuum models

The simplest way to take into account the influence of solvent is to make the assumption that its major effect is to screen the electrostatic interactions in the solute. This can be accomplished by including the appropriate dielectric constant in the denominator of the molecular mechanics term. The calculation is then carried out on the isolated molecule. This method does not take into account of any favourable solvent-solute interactions and the ignores the van der Waals component.

The simplest continuum model is the Onsager[9] model in which the solute occupies a fixed spherical cavity of radius a_0 within the solvent field. A dipole in the molecule, induces a dipole in the medium, and the electric field applied to the solvent dipole will in turn react with the molecular dipole, leading to net stabilisation. A more realistic cavity shape is obtained from the van der Waals radii of the atoms of the solute. The cavity surface is divided into a number of small surface elements with a point charge associated with each surface element. This system of point charges represents the polarisation of the solvent, and the magnitude of each surface charge is proportional to the electric field gradient at that point. The total electrostatic potential at each surface element equals the sum of the potential due to the solute and the potential due to the other surface charges, Equation 2.44.

$$\phi(r) = \phi_\rho(r) + \phi_\sigma(r) \qquad \text{Equation 2.44}$$

where $\phi_\rho(r)$ is the potential due to the solute and $\phi_\sigma(r)$ is the potential due to the surface charges. First the cavity surface is determined from the van der Waals radii of the atoms. The fraction of each atom's sphere which contributes to the cavity is then divided into a number of small surface elements of calculable surface area. The definition of local polar co-ordinate frame at the centre of each atom's van der Waals sphere allows the use of fixed increments of $\Delta\phi$ and $\Delta\theta$ to give rectangular surface elements, Figure 2.4.

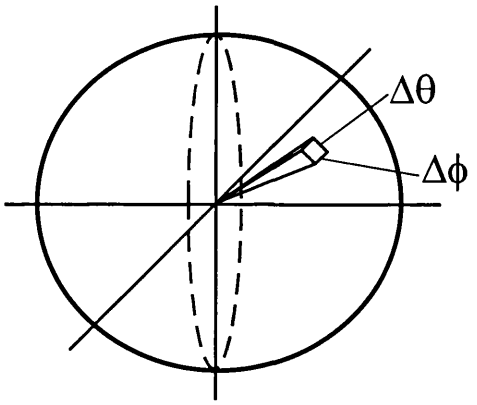
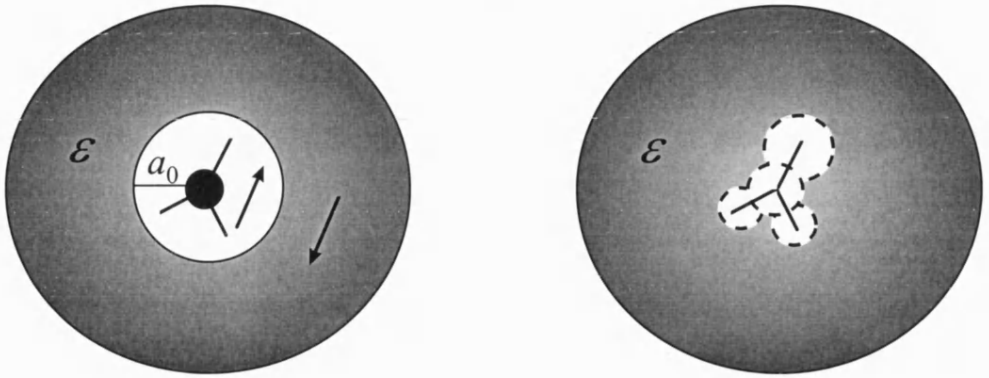


Figure 2.4: Small surface elements can be created on the van der Waals surface of an atom using constant increments of the polar angles, ϕ and θ.

An initial value for the point charge of each surface element is then calculated from the electric gradient due to the solute alone, Equation 2.45.

$$q_i = -\left[\frac{\varepsilon - 1}{4\pi\varepsilon}\right] E_i \Delta S \qquad \text{Equation 2.45}$$

ε is the dielectric constant of the medium, E_i is the electric field gradient and ΔS is the area of the surface element. The contribution of $\phi_\alpha(r)$ due to the other point charges can be calculated by Coulomb's law. These changes are modified iteratively until they are self-consistent. Tomasi's[10] model is an example of a polarised continuum which defines the cavity as the union of a series of interlocking atomic spheres, Figure 2.5.



(a) (b)

Figure 2.5: Diagrammatic representation of (a) Onsager's and (b) Tomasi's continuum models.

2.4.2.1 COSMO

The COSMO[11] dielectric continuum model in MOPAC93[12] was used to mimic solvation. COSMO treats the electrostatics of the solvent as a conductor-like screening model where the solute is embedded in a dielectric continuum of the permittivity, ε, and forms a cavity within the dielectric. The interface between the cavity and the dielectric is known as the 'solvent accessible surface' or SAS. The response of a homogenous dielectric continuum to any charge distribution of the solute consists of the surface charge distribution on the interface. The calculation of the screening charge densities $\sigma(r)$ are implicitly given by Equation 2.46.

$$4\pi\varepsilon\sigma(r) = (\varepsilon - 1)n(r)E^{-}(r) \qquad \text{Equation 2.46}$$

where $n(r)$ is the surface normal vector at a point r and $E^{-}(r)$ denotes the total electronic field on the inner side of the surface at this point. To solve the equation for any shape, segmentation of the SAS into small segments is required. The first step is the construction and segmentation of the SAS. The geometry of the solvent molecules may be described by an effective radius, R^{solv}. This allows the centres of the solvent molecules to be excluded from a sphere of radius $R_\alpha = R_\alpha^{vdW} + R^{solv}$ around the atom α

of the solute with R_α^{vdW} being the van der Waals radius of atom α, Figure 2.6. The R^{solv} term is used in the calculation of R_α to avoid any sharp sphere cut-offs. The effective charges which are responsible for the dielectric are not located at the centres of the solvent molecules. Instead, they are set at a distance δ^{sc} outside the centre, so that their minimum distance to the solute α is $R_\alpha^* = R_\alpha - \delta^{sc}$.

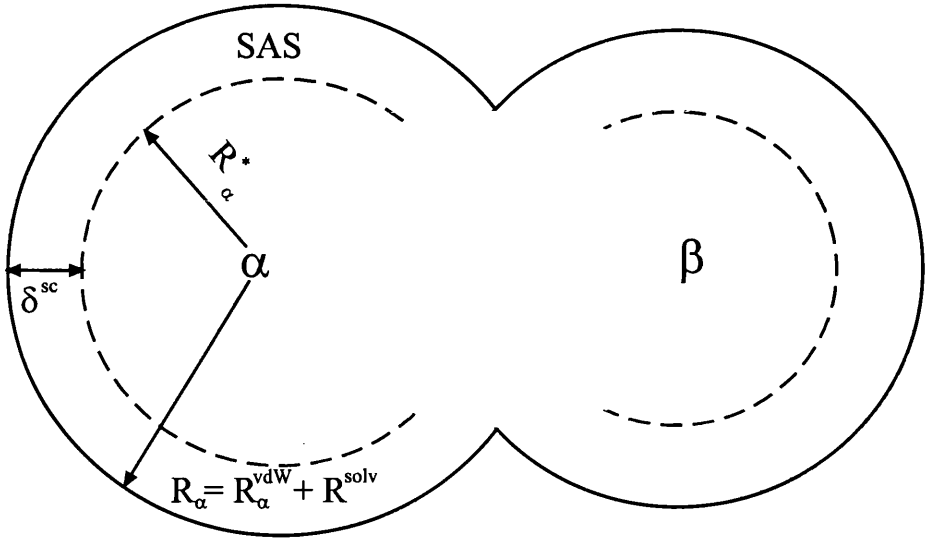


Figure 2.6: Schematic of the construction of the SAS. The solid circles indicate the surface accessible area to the centres of the solvent molecules. The dashed lines indicate the surface accessible to the solvent charges.

The initial preparation of the basic grid of points on the unit sphere is generated by an iterative refinement of triangles starting from a regular icosahedron (a solid figure with 20 faces) generating a large number of points. To ensure that each segment of the SAS is connected to a single solute atom, the basic grid is projected to the R_α sphere and the points lying within the R_β sphere are deleted. The remainder of the points are contracted to the R_α^* sphere. These points represent the allowed positions for the screening charges and define the SAS. The basic points on the SAS are then gathered into segments. The number of segments per atom (NSPA) can be specified. If the NSPA is set to 12, it will correspond to 12 faces of a dodecahedron (the face centres forming an icosahedron) while if the NSPA is set to 32, it will correspond to the 32 faces of a football. For each segment, its area can be calculated for the number of basic points associated with it.

2.5 Transition-state theory

For a single-step reaction, the reactants must pass through a critical structure known as the transition-state which is the point of highest energy along the reaction co-ordinate. Transition-state theory rests on a key number of assumptions. The primary assumption of transition-state theory is that for any elementary process the transition state is in thermodynamic equilibrium with the reactants. The forward reaction, may be expressed as:

$$A + B \rightleftharpoons [TS] \longrightarrow \text{Products} \qquad \text{Equation 2.47}$$

where the equilibrium constant for the transition-state formation $K^{\ddagger}$, by analogy with a regular equilibrium constant, is given by

$$K^{\ddagger} = [TS]/[A][B] \qquad \text{Equation 2.48}$$

A second important element of transition-state theory that is derived from statistical mechanics, is that the universal rate constant from the decomposition of the transition state $k^{\ddagger}$ is given by:

$$k^{\ddagger} = \boldsymbol{k}\, T/h \qquad \text{Equation 2.49}$$

where h is Plank's constant, $\boldsymbol{k}$ is the Boltzmann's constant, and T is the temperature. Since the rate of reaction in terms of reactant concentrations is given by

$$\text{Rate} = k\,[A][B] \qquad \text{Equation 2.50}$$

where k is the rate constant for the reaction, and in terms of the transition-state, the rate constant is given by:

$$\text{Rate} = k^{\ddagger}\,[TS] \qquad \text{Equation 2.51}$$

Equation 2.50 and Equation 2.51 combine to give:

$$k = k^{\ddagger}\,[TS]/[A][B] \qquad \text{Equation 2.52}$$

If $K^{\ddagger}$ from Equation 2.48 is substituted, we get:

$$k = k^{\ddagger} K^{\ddagger} \qquad \text{Equation 2.53}$$

Substituting $k^{\ddagger}$ from Equation 2.49 and rearranging we obtain:

$$K^{\ddagger} = kh/\mathbf{k}T \qquad \text{Equation 2.54}$$

Equation 2.54 gives the relationship between the equilibrium constant for the transition-state formation $K^{\ddagger}$ and that for the rate constant for the reaction. The assumption made here is that the transition-state is in equilibrium with the reactants even though the transition-state is not a 'regular' species.

2.6 Minimisation algorithms

Minimisation algorithms find the nearest local minimum to any given conformation.[13] These algorithms usually do not find the global minimum of the system. All minimisation algorithms stem from the fact that, in general, any function F(x) of a variable x can be expanded in a Taylor series around the minimum x_0,

$$F(x) = F(x_0) + (x-x_0)F^{|}(x_0) + \tfrac{1}{2}(x-x_0)^2F^{||}(x_0) + \ldots \qquad \text{Equation 2.55}$$

where $F^{|}$ and $F^{||}$ are the first and second derivatives of the function. The extension to a multi-dimension function requires replacing the variable x by a vector $|x|$ and the derivatives by the appropriate gradients. Different minimisation algorithms are classed according to their order, i.e. according to the highest derivative used by the algorithm. hence, order 1 algorithms use only the slope information while order 2 also take into account information about the curvature of the surface.

2.6.1 Order 0 algorithms

Order 0 methods do not directly use information about the slope and curvature of the surface as it minimises. As a result, these methods are very crude but work well for surfaces that are simple. A grid search is an example of an order 0 minimisation algorithm. In this method a regular grid is placed upon the surface and the value of the function at each node is calculated. The grid point with the lowest energy is chosen as the minimum.

2.6.2 Order 1 algorithms

Order 1 algorithms use the first derivative of the function (the gradient of the multi-dimensional energy surface) to direct the search towards the nearest local minimum. This means that they use information about the slope but not about the curvature (which is given by the second derivative). The algorithms try to compensate for this lack of information by using different 'tricks'. The steepest descent algorithm uses a variable step size for this purpose while the conjugate gradients algorithm[14] makes a sophisticated use of the previous steps for a similar goal. In general these methods iterate over the following equation in order to perform the minimisation.

$$R(k) = R_k\text{-}1 + 1k\ S_k \qquad \text{Equation 2.56}$$

where R_k is the new position at step k, R_k-1 is the position at the previous step k-1, $1k$ is the size if the step to be taken at step k and S_k is the direction of that step. The different methods differ in the way they choose the step size and the step direction.

2.6.2.1 Steepest descent

At each step of the steepest descent algorithm, the gradient of the potential *gk* (i.e. the first derivative in multi-dimensions) is calculated and a displacement is added to all the co-ordinates in a direction opposite to the gradient (i.e. in the direction of the force). In terms of the general scheme outlined in Equation 2.57, this means that at

$$Sk = -gk \qquad \text{Equation 2.57}$$

At every iteration, the step size, *1k*, is adjusted to compensate for the lack of curvature information. If the energy drops, the step size is increased by 20% to accelerate the convergence. If the energy rises due to passing over the minimum, the step size is halved. Because of the finite step sizes, the method does not follow the gradient smoothly down to the minimum, but rather oscillates about it and as a result it usually gets stuck around a minima.

2.6.2.2 Conjugate gradients

A conjugate gradients algorithm uses information from the current gradient and that from previous steps. By using the previous gradients the method compensates for the lack of curvature information. Namely, while the first step is taken along the gradient, i.e.

$$S1 = -g1 \qquad \text{Equation 2.58}$$

For all steps $k > 1$, the direction of the step is a weighted average of the current gradient and the previous step direction, i.e.

$$Sk = -gk + bk\ Sk\text{-}1 \qquad \text{Equation 2.59}$$

The weight factor *bk* is calculated from the proportion of the magnitude of the current and previous gradients. For quadratic surfaces of dimension n it can be shown that conjugate gradients converges at the nth step. But for non-quadratic surfaces such as molecular surfaces, it converges more efficiently than steepest descent. The disadvantage of the conjugate gradients algorithm is that it can accumulate numerical errors along the way.

2.6.3 Order 2 algorithms

Order 2 algorithms use both the first derivative (i.e. the slope) and the second derivative (i.e. curvature) during the minimisation. This means that for a molecule of N atoms it requires not only the vector of $3N$ first derivatives but also the Hessian matrix of $(3N)^2$ second derivatives.

2.6.3.1 Newton-Raphson

The basic idea behind the Newton-Raphson minimisation algorithm can be written in the one-dimensional case as:

$$xk+1 = xk - F^{I}(xk)/F^{II}(xk) \qquad \text{Equation 2.60}$$

where $xk+1$ is the next position, xk is the current position and F^{I} and F^{II} are the first and second derivatives at the current position. Extending this to a multi-dimensions requires both the gradient vector and the Hessian matrix of the second derivatives. Although the Newton-Raphson algorithm is very accurate and converges well, it is hard to apply to large systems as the need to calculate the Hessian matrix every iteration makes the algorithm computationally expensive.

2.6.3.2 Adopted-Basis-Newton Raphson (ABNR)

The ABNR[15] algorithm is a useful second derivative method which is particularly suited for large systems. Rather than using the full set of vectors as in the Newton Raphson method, ABNR uses a smaller basis that is limited to the subspace in which the system has made the most progress in the past moves. In this way the system moves in the best direction in a restricted subspace. For this small subspace, a second derivative matrix is constructed numerically from the change in the gradient vectors, and is inverted by an eigenvector analysis allowing the routine to recognise and avoid saddle points in the energy surface.

2.7 Saddle searching algorithms

2.7.1 Eigenvector-following algorithm (EF)

Transition-states are characterised as stationary points having one, and only one negative hessian eigenvalue (the hessian being the matrix of second energy derivatives with respect to the geometrical parameters varied) whereas minima possesses all positive eigenvalues.[16] The requirement of a negative hessian eigenvalue means that one has to be much more careful as regards to the type of step taken on the energy surface when searching for a transition-state than a minima. As a consequent, transition-states are much harder to find than a minima. For a transition-state search, if you are in the region of the energy surface where the hessian has the required one negative eigenvalue, then the Newton-Raphson step is a good step to take as it maximises along one mode while minimising along all others. The 'eigenvector-following' algorithm by Baker,[17] which is incorporated in the Gaussian[18] and MOPAC93 packages, fulfils the above criteria. The EF algorithm uses an initial guess for the transition-state and then the gradient vector and the hessian matrix are calculated for it. The hessian is then diagonalised and the number of negative eigenvalues are determined. The negative eigenvalue that corresponds to the correct mode is then followed.

2.7.2 Conjugate peak refinement (CPR)

The conjugate peak refinement saddle searching algorithm attempts to locate saddle points by considering two aspects of their nature. The first aspect is that the saddle point is a maximum on the minimum potential energy 'valley' that connects reactants and products on a potential energy surface[19]. The second aspect is that a set of conjugate vectors describing displacements about the saddle point comprise a stationary maximum in one direction and stationary minima in all other directions; that is, the hessian at the saddle point has one negative eigenvalue.

CPR is initiated with a set of points for structures connecting reactants to products. The vector describing geometric motion between the two points of highest energy is taken as the stationary-maximum conjugate vector. Maximisation along this vector and minimisation along a number of the other conjugate vectors is performed sequentially. A new energy maximum along the resultant set of points is then found and the process

repeated. The line minimisations are performed after the line maximisation and then in tandem. The direction for each line minimisation is generated as being conjugate to the last direction minimised and starting at the point on the potential energy surface that was the result of the previous line minimisation.

If the surface is quadratic, then each minimisation will remain conjugate to the original maximisation direction. Where the surface is non-quadratic, this may not be so. As a consequence, line minimisations will interfere with the maximisation direction (that is become non-conjugate to it) and the algorithm will converge toward a minimum not a saddle point. It is partly possible to overcome this problem, by maximising along one conjugate direction and minimise along all others simultaneously. This should ensure that the minimisations are all conjugate to the maximisation. This is how the partial rational function operator method operates.[20]

2.7.3 Partial rational function operator (P-RFO)

The P-RFO algorithm constructs a step that represents maximisation in the direction of one of eigenvector and minimisation in all others. The repeated application of this approach leads to the maximised direction being at a stationary maximum and all other directions being at stationary minima eventually leading to a saddle point. In CPR, the vector defining that direction which was to be maximised was defined as the translation vector between two points on either side of a maximum. Such an approach can not work for the P-RFO method as only one point is defined. The mechanism for defining the maximised direction is to choose one of the eigenvectors of the hessian that is most similar, in terms of direction, to the eigenvector maximised in the last optimisation step. This concept is called mode following where the name mode represents an eigenvector of the hessian. Maximisation in the direction of a mode should make the eigenvalue associated with that mode more negative (or less positive). If the followed mode is the one with the lowest eigenvalue, then it may be possible to assume that the algorithm can continue to follow that the mode with the lowest associated eigenvalue.

2.8 Intrinsic reaction co-ordinate (IRC)

From the minima and saddle points it is a simple matter to compute the intrinsic reaction co-ordinate.[21] The IRC is the path that would be followed by a particle moving along the steepest descents path with an indefinitely small step from the transition structure down to each minima when the system is described using mass-weighted co-ordinates. The initial directions towards each minima can be obtained directly from the eigenvector that corresponds to the imaginary frequency at the transition-state structure. A simple steepest descents algorithm with a reasonable step size will usually give a path that oscillates about the true minimum energy path, Figure 2.7. The objective of the IRC algorithm is to locate the minimum as efficiently as possible.

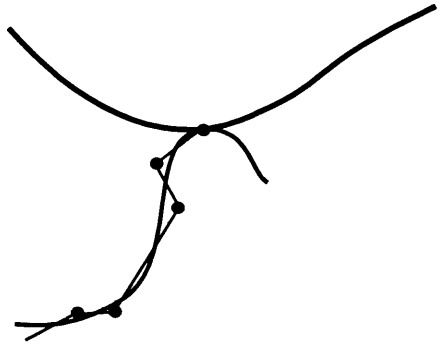


Figure 2.7: A steepest descents minimisation algorithm produces a path that oscillates about the true reaction pathway from the transition structure to a minima.

Chapter Three

Reactivity Probes

Chapter 3- Reactivity Probes

3.1 Kinetic isotope effects

The kinetic isotope effect of a reaction is a widely used tool for elucidating mechanisms. The most common isotopic substitution is a deuterium for a hydrogen atom, although isotope effects for heavier atoms are frequently used. Substitution of one isotope for another does not alter the shape of the potential energy functions describing molecular reactivity, but does affect rate and equilibrium constants. The comparison of calculated isotope effects with experimental isotope effects on the reaction rate constant can provide detailed information of bonding changes occurring during the chemical reaction in solution.

3.1.1 Origin

Using transition-state theory and statistical mechanics it is possible to derive an expression which relates the kinetic isotope effect (KIE) to a ratio of complete partition functions, Equation 3.1.[22]

$$k_{H/D} = (Q_D/Q_H)/(Q_D^{\ddagger}/Q_H^{\ddagger}) \qquad \text{Equation 3.1}$$

The KIE of a reaction has its origin in the effect on the molecular mass, the moments of inertia, molecular vibrations of the substrate and of the corresponding vibrations of the transition-state, brought about by the isotopic substitution. The complete partition function maybe expressed as a product in which the separate terms represent the partition functions for the various degrees of freedom, Equation 3.2.

$$Q = q_{trans} \cdot q_{rot} \cdot q_{vib} \qquad \text{Equation 3.2}$$

The substitution of deuterium for hydrogen will have a minor effect on the molecular masses and the moments of inertia. The transitional and rotational partition functions are only slightly affected. Therefore the most important factor is the effect that isotopic substitution has on the vibrational partition function. The KIE arises from the fact that the zero-point energy of a molecule containing a heavier isotope is lower than that of the lighter, unsubstituted molecule. Consequently, less energy is required to dissociate a carbon-hydrogen bond than a carbon-deuterium bond. If a C-H bond is cleaved in the rate determining step, substitution of a deuterium for a hydrogen atom will result in a rate decrease. Therefore the ratio of the rate constants for the unsubstituted and the substituted reaction will be greater than unity.

3.1.2 Secondary KIE

Chemical rates of reaction and equilibrium constants change with isotopic substitution. Where the substitution is on an atom which is not bond making or breaking, these are called secondary isotope effects. Secondary KIEs arise from changes in the zero-point energies in the vibrations of the isotopically substituted bonds, Figure 3.1. The smaller vibrational frequencies in the transition-state imply smaller zero-point energies for the vibrations, compared to the reactant. Since a C-D bond has lower zero-point energy than a C-H, less zero-point energy is lost at the transition-state for the deuterated compound than for the corresponding hydrogen compound. The total energy difference between the ground-state and the transition-state is therefore greater for the deuterated compound which in turn reacts more slowly.

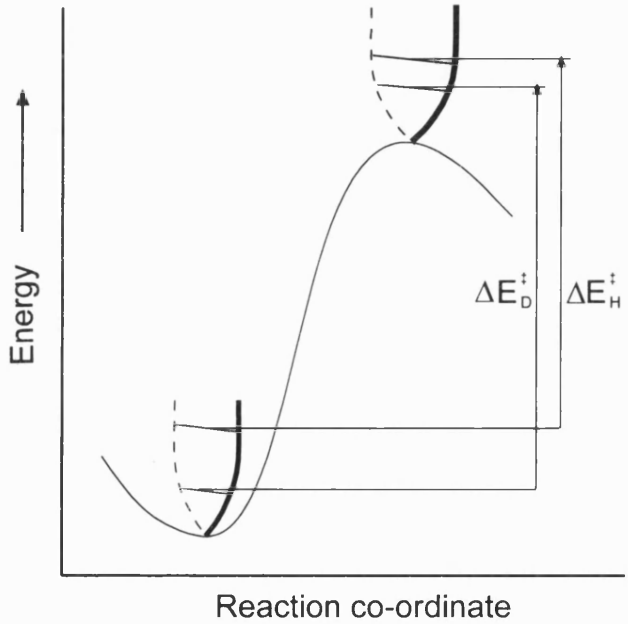


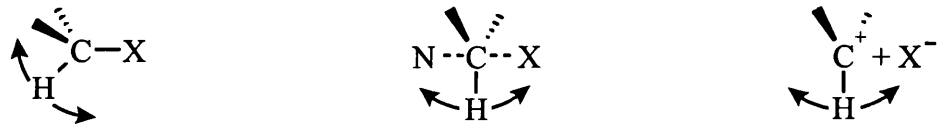
Figure 3.1: A representation of the zero-point energies responsible for a secondary deuterium isotope effect.

$R–X + N^- \rightarrow [\ X–R–N]^{-\ddagger} \rightarrow R–N + X^-$	S_N2	Equation 3.3
$R–X \rightarrow R^+ + X^-$	S_N1	Equation 3.4
$R^+ + N^- \rightarrow R–N$		Equation 3.5

Secondary KIE are most commonly observed when a deuterium substitution is made at a carbon which undergoes a change in hybridisation during the reaction. If the bond is exchanged in one act, the substitution is defined as bimolecular nucleophilic substitution (S_N2), Equation 3.3. Alternatively, if the bond to the leaving group is

broken and reconstituted in separate acts, the substitution is defined as unimolecular nucleophilic substitution (S_N1), Equation 3.4. Values of k_H/k_D close to unity have been attributed to S_N2 mechanisms, and values between 1.10 and 1.25 have been attributed to the S_N1 mechanism. This has in turn has been used to describe the tightness or looseness of the transition-state.

The differences between isotopic values have been rationalised in terms of the vibrational characteristics of the ground-state and those of different types of transition-states, Figure 3.2. The vibration of importance is the HCX bending mode in the ground-state as compared to the out-of-plane bending in the transition-state. The ground-state species are sp^3 hybridised while at the transition-states, they are sp^2 hybridised. This manifests itself in a significant decrease in the force constant for the out-of-plane bending. The energy difference between the zero-point energies is less in the transition-state therefore, the reaction is slowed by the substitution of the deuterium for the hydrogen atom. For an S_N2 transition-state, the partial bonding of both the nucleophile (N) and leaving group (X) impedes the bending of the hydrogen, and so there will be little or no overall change in the ease of vibration as compared to the ground-state. For an S_N1 transition-state, the breaking of the C-X and the change in hybridisation leads to a greater freedom of motion for the bending of the hydrogen than is found in the ground-state.



(a) (b) (c)

Figure 3.2: A schematic showing the α–hydrogen bending motion associated with (a) the ground-state sp^3 hybridised and (b) S_N2 and (c) S_N1 transition-states that are sp^2 hybridised.

It has been shown computationally that secondary α-deuterium KIEs arise from a varying combination of the bending force constants of HCX and the C-H stretching force constants. AM1 calculated α-deuterium kinetic isotope effects for the degenerate reaction, Equation 3.6,

$$Cl^- + CHR^IR^{II}Cl \rightarrow CHR^IR^{II}Cl + Cl^- \qquad \text{Equation 3.6}$$

where R^I = H or Me, and R^{II} = H, Me or MeO shows that the KIE is dominated by the zero point energy. The C-H stretch contribution to the zero point energy is almost a

constant inverse contribution with the other modes varying proportionally with k_H/k_D. The KIE for methyl transfers are inverse in nature.[23,24] The KIE arises from an increase in the C-H stretching force constant accompanied by a change from tetrahedral to trigonal geometry about the α-carbon along the reaction co-ordinate and not from the bending force constant of HCX.[25] The overall KIE's for the degenerate S_N2 reactions increased as the transition-state became looser. Subsequent *ab initio* studies on S_N2 transition-states and their relationship to KIE concurred with the above conclusions.[26,27]

3.1.3 Modelling secondary KIEs

Secondary isotope effects can be modelled by approximating the vibrations to harmonic oscillators, the nature of which can then be determined by force constants.

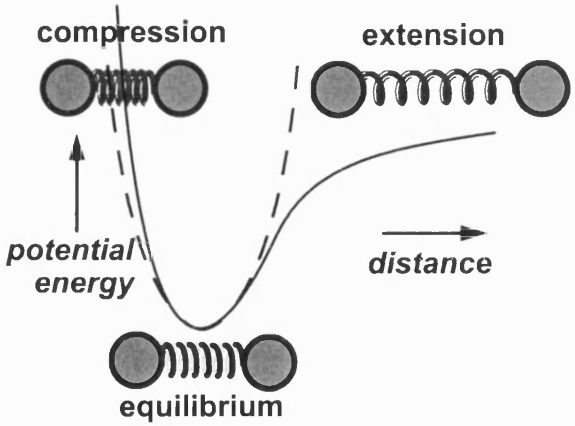


Figure 3.3: The compression or extension of a spring showing the increase or decrease in potential energy. The function describing how the energy varied with the inter-atomic distance can be approximated to a parabolic curve shown as a dashed line.

A diatomic molecule such as H_2 vibrates. The bond may be regarded rather like a spring. The compression or extension of the spring away from its equilibrium length involves raising the potential energy of the molecule. If the spring obeyed Hooke's law, the increase in potential energy would be proportional to the square of the displacement, and the function describing how the energy varied with the inter-atomic distance would be a parabolic curve like the dashed line in Figure 3.3. Real molecules are harder to compress than would be predicted by a parabolic energy curve due to the short-range repulsive interactions between the atoms which make the energy rise more steeply as the inter-atomic distance is decreased. On the other hand, real molecules are easier to stretch than would be predicted for a harmonic oscillator obeying Hooke's law as the

extension of the bond in the H_2 molecule leads eventually to dissociation, with the potential energy curve flattening out as the inter-atomic distance is increased. The shape of the actual potential energy profile is more closely approximated by a Morse function, Equation 3.7

$$E = D_e (1 - \exp(-\beta (l_o - l_i)))^2 \qquad \text{Equation 3.7}$$

where D_e is the dissociation energy, l_i is the equilibrium bond length and l_0 is the studied bond length. β is a coefficient describing the curvature of the function, given by $\beta = (k_e/2D_e)^{1/2}$ where k_e is the Hooke's law, harmonic oscillator force constant at the energy minimum.

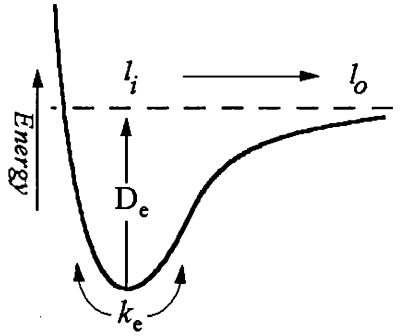


Figure 3.4: A Morse curve where l_i is the equilibrium bond length at the minimum of the energy curve and D_e is the bond dissociation energy as measured from the energy minimum.

The parabolic Morse curve, Figure 3.4, with force constant k_e is a good approximation to both the Morse curve and the actual potential energy profile for *small* displacements away from the equilibrium bond length. The frequency ω with which the molecule vibrates about its equilibrium position may therefore be described approximately by the harmonic oscillator expression, Equation 3.8, where μ is the reduced mass of the diatomic molecule A–B, given by $1/\mu = 1/m_A + 1/m_B$, and has dimensions of reciprocal time (units of s^{-1}).

$$\omega = (1/2\pi) (k_e / \mu)^{1/2} \qquad \text{Equation 3.8}$$

It is common, however, to express vibrational frequencies ν in units of cm^{-1}, where $\nu = \omega/c$ (velocity of light).

3.1.4 CAMVIB

The diagonalisation of the mass-weighted Cartesian force constant matrix evaluated at a stationary point on the potential energy surface should yield six zero eigenvalues for translational and rotational motions and 3N-6 non-zero eigenvalues for vibrational modes. In practice, Cartesian calculated force constants do not lead to exactly zero eigenvalues for the translational and rotational motions. The unwanted contamination by spurious translational and rotational contributions give rise to small non-zero frequencies for the translational and rotational motion. The non-zero frequencies need to be removed before the kinetic isotope effect is calculated and this is achieved by the CAMVIB program.[28] The small non-zero translational and rotational frequencies arising from force-constants that are computed directly in Cartesian co-ordinates are eliminated by means of a projection method which preserves only the internal co-ordinates contribution to the force constants. The resultant vibrational frequencies are then said to be 'pure' as there are six zero eigenvalues for translational and rotational motions and 3N-6 non-zero eigenvalues for the vibrational modes.

3.1.5 CAMISO

CAMISO calculates the non-potential energy contributions to thermodynamic properties for molecules or transition-states. A free energy change may be expressed by Equation 3.9 in which Δ_ε, ΔE_{zp}, ΔE_{th}, and $P\Delta V$ represent the change in potential energy, zero-point energy, thermal energy, and the pressure-volume terms, respectively.

$$\Delta G = \Delta_\varepsilon + \Delta E_{zp} + \Delta E_{th} + P\Delta V - T\Delta S \qquad \text{Equation 3.9}$$

The zero-point energy term is given by simply by changes in vibrational frequencies, the thermal energy and temperature-entropy term are determined by partition functions q' for the reactants and q'' for the transition-state as given by Equation 3.10 and Equation 3.11.

$$\Delta E_{th} = RT^2 \, d \ln (q'' / q') / dt \qquad \text{Equation 3.10}$$

$$T\Delta S = RT \ln (q'' / q') \qquad \text{Equation 3.11}$$

For any molecular species the partition functions for translational, rotational, vibrational, and free energy of rotation motions can be calculated using standard expressions of thermodynamics within the rigid rotor/harmonic oscillator approximation to obtain thermodynamic functions including (G-ε), E_{zp}, (H-ε) and S.

Input of a molecular force-field into CAMISO simply requires the specification of individual internal co-ordinates and their associated force-constants which is supplied by the CAMVIB program. A model force-field for the calculation of vibrational partition functions by the standard harmonic oscillator expression, Equation 3.12, in which ν_i is the frequency (in cm^{-1}) of the ith normal mode.

$$q_{vib} = \prod_i [1 - \exp(-hc\nu_i) / kT]^{-1} \qquad \text{Equation 3.12}$$

The potential energy V is given by

$$V = \frac{1}{2}\sum_i F_{ii}(\Delta R_i)^2 \qquad \text{Equation 3.13}$$

where the summation is over changes in all internal co-ordinates and ΔR_i and F_{ii} are the diagonal force-constants.

3.2 Structure-activity relationships

The experimental knowledge of the transition-state largely derives from observations of the changes in reactivity as a function of the variation of some structural parameter in the molecule. The magnitude of the effect depends on the coupling between the structural change and the site of the reaction. If the parameter is remote from the reactive site, its effect is reduced. The application of structural effects is carried out by correlating the free energy of the reaction under investigation against that of a standard process. Such correlations are usually known as linear free energy relationships.[29] Linear free energy relationships provide a means of predicting unknown empirical data and are a most valuable source of mechanistic information.

The linear correlation between the logarithms of rate constants and pK_a values of substituted benzoic acids, known as the Hammett relationships, have been used as a means of estimating the relative polarity of substituents. The measure of polarity, the Hammett σ-value for the substituent, is simply the effect of that substituent on the dissociation constant of the substituted benzoic acid, (see Chapters 5 and 6 for more detail). A plot of the logarithm of the rate constant for a proton transfer reaction against pK_a of the acid acting as a variable proton donor or against the pK_a of the conjugate acid of the base acting as a variable proton acceptor is linear in nature which is known as a

Brønsted correlation. The slopes of these plots are given by the symbols α and β, for the general acid and base catalysis, respectively, (see Chapter 7 for more detail).

Both the Hammett and Brønsted linear free energy relationships have been used to determine a change in mechanism or a change in rate determining step, respectively. There is said to be a change in mechanism if the observed rate constant on one side of the breakpoint is greater than that calculated from the correlation on the other side. The Hammett correlation vs. σ^+ for the acetolysis of substituted benzyl *p*-toluene sulphonates, Figure 3.5(a), shows two linear correlations. From the interpretation of the Hammett plot the reaction is said to occur by the S_N2 mechanism when the substituent effect is electron-withdrawing and as the substituent is made more electron-donating, the S_N1 transition-state is stabilised and there is a changeover in mechanism.[30] A change in the rate-determining step is observed if the rate constant on one side of the breakpoint is less than that calculated from the correlation on the other side, Figure 3.5(b).

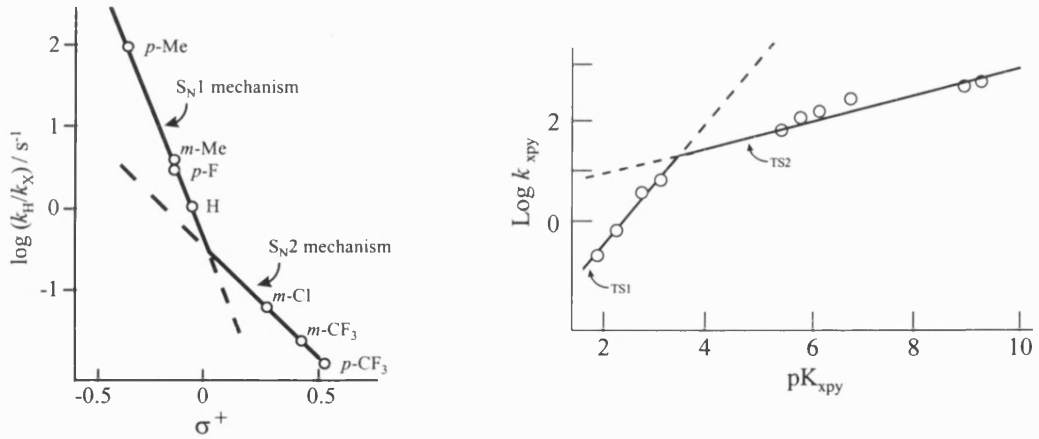


Figure 3.5: (a). A Hammett correlation, solid lines, showing a change in mechanism. (b). A Brønsted correlation, solid lines, showing a change in the rate-determining step.

A Brønsted plot can be seen for the attack of various substituted pyridines on 2,4,6-trinitrophenyl acetate.[31] The plot comprised of two straight lines that intersected. The presence of this break point between two linear portions of a free energy relationship is regarded as evidence for the existence of the formation or breakdown of an intermediate in a stepwise mechanism. In both examples, only the solid lines are seen experimentally, whereas both the solid and dashed lines could be calculated computationally.

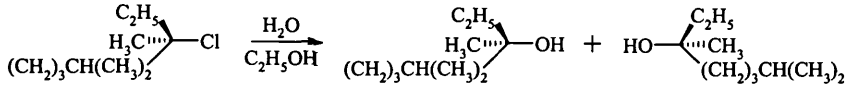
Chapter Four

A Theoretical Investigation into Nucleophilic Substitution at a Saturated Carbon Centre

Chapter 4- A Theoretical Investigation into Nucleophilic Substitution at a Saturated Carbon Centre

4.1 Introduction

Aliphatic nucleophilic substitution reactions involve the replacement of one functional group by another. The mechanism of nucleophilic aliphatic substitution was one of the first reactions to be studied in depth. During the 1930's much of the mechanistic background knowledge was assembled and interpreted by Ingold and Hughes.[32] It was proposed that nucleophilic substitution reactions could proceed by one of two different mechanisms; the bimolecular nucleophilic substitution (S_N2) or the unimolecular nucleophilic substitution (S_N1) mechanisms. The S_N2 reaction was stated to occur with complete inversion of stereochemistry at a tetrahedral carbon centre and to show second-order kinetics. The S_N1 reaction was stated to occur with complete racemisation and show first-order kinetics. The expectation that S_N1 reactions on chiral reactants lead to racemic products has been borne out by experiment. However only a few reactions occur with complete racemisation. For example, the solvolysis reaction of optically active (*R*)-6-chloro-2,6-dimethyloctane leads to 40% (*R*) and 60% (*S*), Equation 4.1. [33]



Equation 4.1

(*R*)-6-Chloro-2,6-dimethyloctane 40% *R* 60% *S*

There are four possible degrees of alkyl substitution on a saturated carbon which are referred to as tertiary, secondary, primary or methyl compounds. The reactions of methyl or primary compounds are said to react by the S_N2 mechanism while tertiary compounds react by the S_N1 mechanism. This viewpoint is still reported in every introductory organic textbook. However, the position of secondary compounds are usually skipped over or even overlooked! Do they react by an S_N2 or S_N1 mechanism or a mixture of both? More advanced organic texts have attempted to address the situation of nucleophilic aliphatic substitution, however no concise, non-contradicting picture of nucleophilic substitution has emerged.

4.2 Mechanistic definitions

4.2.1 Ingold and Hughes

Ingold and Hughes were the first to use the term bimolecular nucleophilic substitution, S_N2, and unimolecular nucleophilic substitution, S_N1, mechanisms to describe a reaction. They stated that if the bond is exchanged in one act, the substitution is defined as S_N2. Alternatively, if the bond to the leaving group is broken and reconstituted in separate acts, the substitution is defined as S_N1. Which of the mechanisms operates depends upon the structure of the reactants and the type of solvent used. In a series of substituted alkanes such as

Methyl, Ethyl, ^{i}Propyl, ^{t}Butyl,

the alkyl group will steadily increase its inductive power until a changeover in mechanism occurs.[34] At first, the S_N2 mechanism would prevail up to a certain point after which the S_N1 mechanism would take control. Accordingly, before the transition point the substitution should exhibit second order kinetics with a decreasing rate and post-transition point, the substitution should exhibit first order kinetics with an increasing rate. They proposed that at this transitional point both the S_N2 and S_N1 mechanisms would be present.[35] Experimentally they showed that the changeover in mechanism for the solvolysis reaction for a series of alkyl bromides in 60% aqueous ethanol occurred between the ethyl (S_N2) and *iso*-propyl (S_N1) bromide.

4.2.2 Winstein - the ion pair mechanism

Winstein proposed that a reaction was 'nucleophilic' (N) if there was a covalent interaction between the nucleophile and substrate in the rate determining transition step. If no interaction existed in transition-state with the nucleophile then it would be classed as 'limiting' (Lim).[36] The transition-states of aliphatic substitutions are considered to be a mesomeric form of the canonical structures in Equation 4.2, with the covalent interaction being represented by giving weight to canonical 3.

$$\underset{1}{N:^{-}\quad R\text{-}X} \qquad \underset{2}{N:^{-}\quad R^{+}\quad {}^{-}:X} \qquad \underset{3}{N\text{-}R\quad {}^{-}:X} \qquad \text{Equation 4.2}$$

The difference between the Ingoldian classification and that due to Winstein was that Ingold proposed two separate mechanisms while Winstein regarded there to be one, merging mechanism. These differences only become important when discussing borderline mechanisms.

4.2.3 Bentley and Schleyer - the S_N2(intermediate)

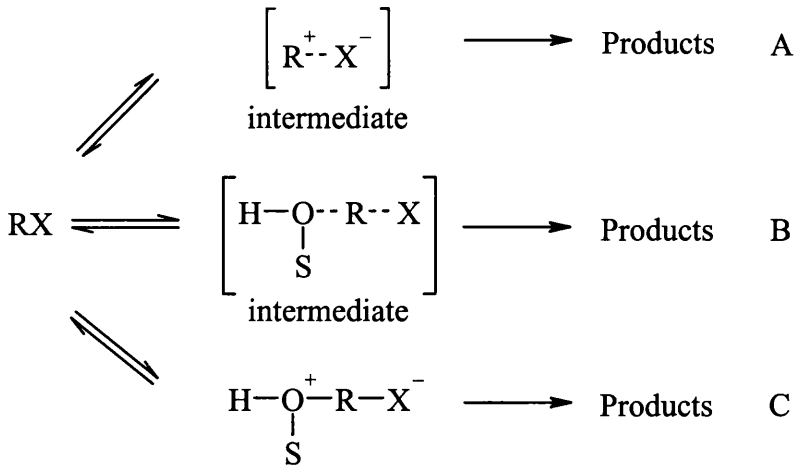


Figure 4.1: Spectrum of solvolysis mechanisms as described by Bentley and Schleyer. (A) represents the S_N1 transition-state leading to intermediate intimate ion pair that is not nucleophilically solvated. (B) represents the S_N2(intermediate) nucleophilically solvated transition-state leading to a nucleophilically solvated ion pair intermediate. (C) represents the S_N2 transition-state. The solvation by hydrogen bonding to the leaving group is not shown.

Instead of an ion pair intermediate playing a role, Bentley and Schleyer believed that a varying degree of nucleophilic solvent assistance existed.[37] They defined this assistance as a kinetically significant involvement of the solvent or nucleophile by partially bonding to any atom of the substrate as distinct from general electrostatic solvation. Nucleophilic solvent assistance may be described as the 'S_N2 character' of the system.[38] In their view, solvolysis reactions exhibited varying degrees of this nucleophilic assistance ranging from very strong (classical S_N2) to negligible (classical S_N1), Figure 4.1. In the borderline region a third mechanism, the S_N2(intermediate) is said to operate.[39] The S_N2(intermediate) mechanism involves a weak nucleophilic assistance by the nucleophile in which the solvated nucleophile-leaving group ion pair corresponds to an energy minimum, i.e. an intermediate. Nucleophilic assistance increases from the S_N1 to S_N2(intermediate) to S_N2 mechanisms. The increase in nucleophilic assistance reduces the activation energy for heterolysis for the S_N2(intermediate) and S_N2 mechanisms and decreases the relative stability of the intermediate, Figure 4.2.

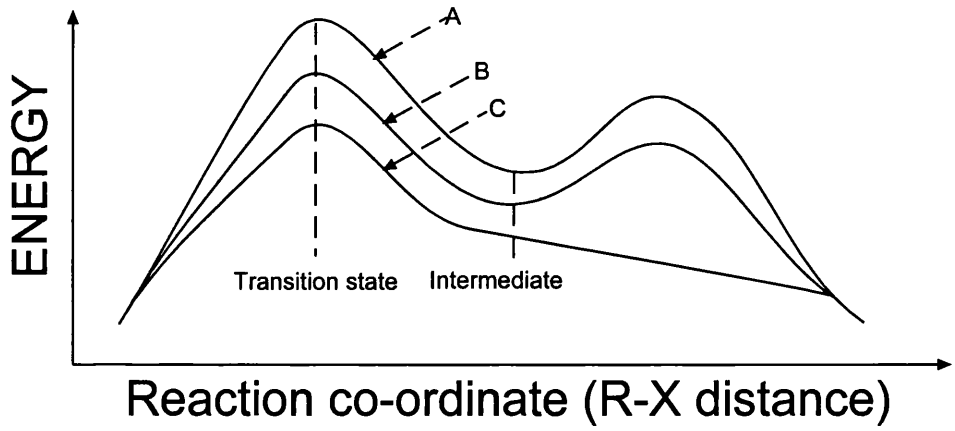
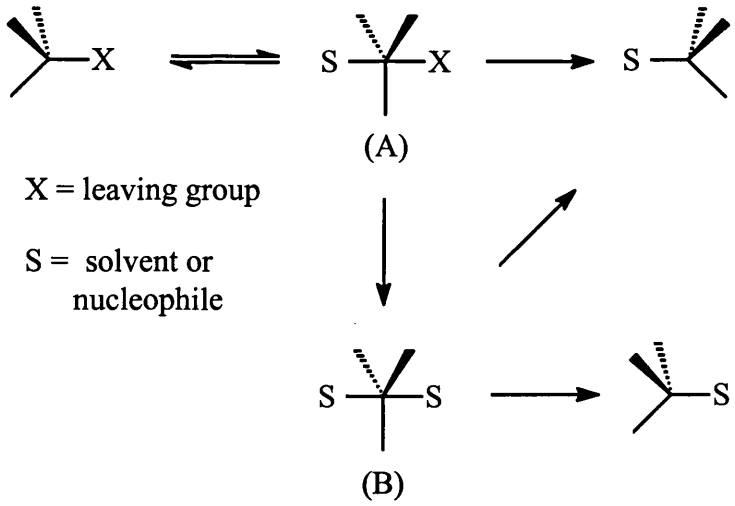


Figure 4.2: A schematic representation of the upper portion of potential energy surface for the merging of mechanisms. Profile A is for the S_N1 mechanism, profile B is for the S_N2(intermediate)mechanism and profile C is for the S_N2 mechanism.

4.2.4 Doering and Zeiss - the structural hypothesis

The gradual merging of nucleophilic substitution mechanisms was also developed by Doering and Zeiss which they termed 'the structural hypothesis'.[40] In this explanation of nucleophilic substitution, explicit account is taken of the role of stabilisation of the carbocation by the solvent, nucleophile or leaving group. The simplistic assumption is made that only two interaction sites are needed to be considered, Equation 4.3.



Equation 4.3

The rate limiting step is the formation of the intermediate (A) in which the nucleophile or solvent and leaving group provide the stabilisation. This intermediate may collapse to give starting material or products with an inversion of configuration. Alternatively, the leaving group may be replaced by a solvent molecule or nucleophile (B) which could give rise to a racemic product. The stability of the intermediate is vital to which reaction path would be followed.

The various possibilities are illustrated in Figure 4.3. For a given nucleophile and leaving group, the stability of (A) will parallel the stability of the carbocation obtained on ionisation. In a case where ionisation would be energetically unfavourable, so that (A) could only be formed by nucleophilic assistance, it could be thought of as a process that resembles the S_N2 mechanism and (A) would be a transition-state. As (A) becomes more stable it is possible to regard it as an intermediate marked by the appearance of a shallow minimum in the potential energy curve. In the limiting case there is no interaction of the nucleophile in the rate determining step and this can be classed as an S_N1 mechanism. Doering and Zeiss' structural hypothesis classed the S_N2 and S_N1 mechanism as extremes of a single, merging mechanism.

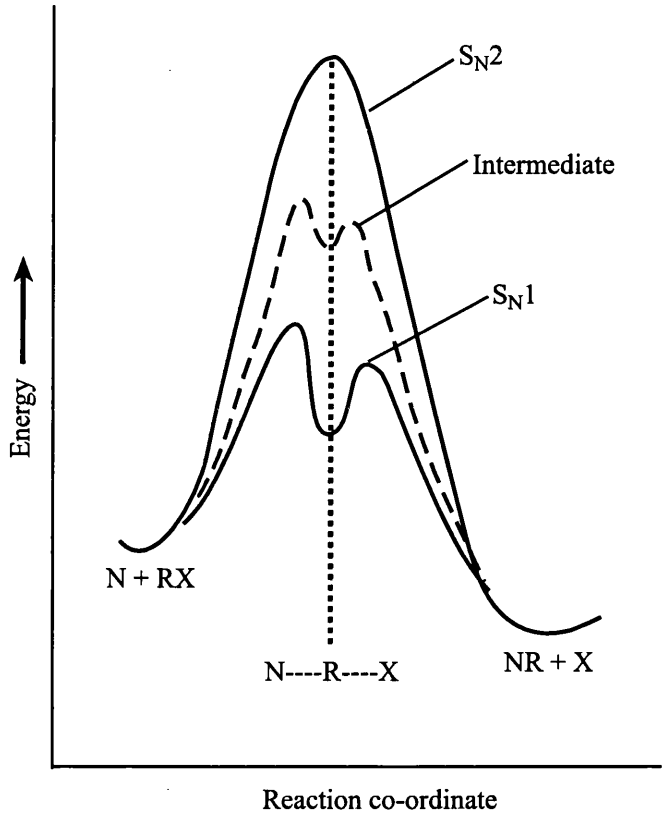


Figure 4.3: Potential energy curves for different nucleophilic substitutions.

4.2.5 Sneen

According to Sneen, all S_N1 and S_N2 reactions form one mechanism, that of the ion pair mechanism.[41] In this mechanistic description, the substrate first ionises to an intermediate ion pair which is then converted to products, Equation 4.4.

$$RX \underset{}{\overset{k_1}{\rightleftharpoons}} R^+X^- \overset{k_2}{\rightleftharpoons} \text{Products} \qquad \text{Equation 4.4}$$

The difference between the S_N1 and S_N2 mechanism is that for the rate of formation for the S_N1 ion pair, k_1 is rate determining while for the S_N2 mechanism, k_2 is rate determining, Figure 4.4. The borderline behaviour of a mechanism will exist when the magnitude of k_1 and k_2 are equal to each other.[42]

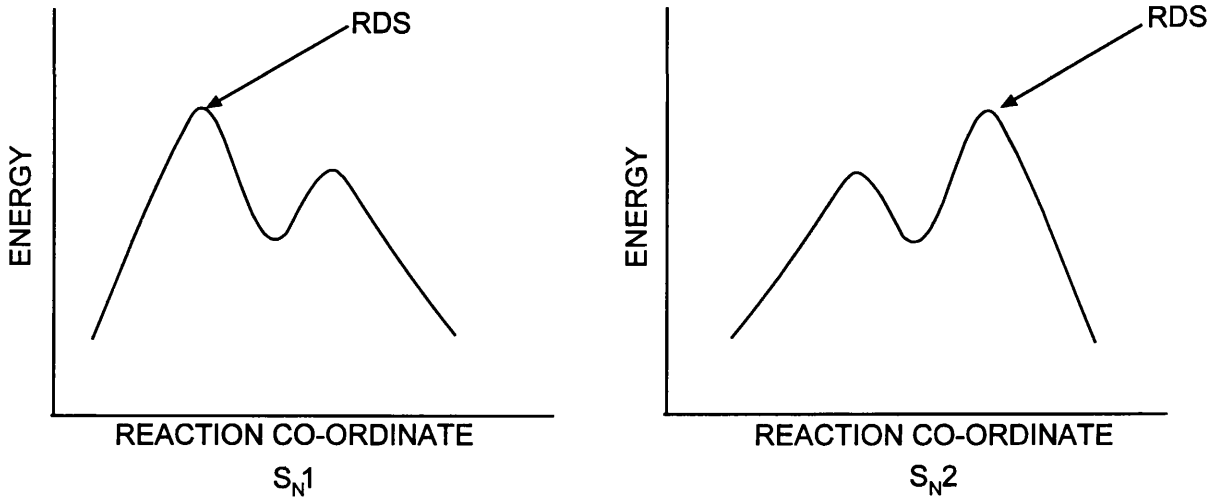


Figure 4.4: Potential energy curves depicting the different rate determining steps (RDS) for the S_N1 and S_N2 nucleophilic substitutions according to Sneen.

4.3 Energy surfaces and the S_N1/ S_N2 mechanistic borderline

It is possible to construct a 2-dimensional 'map of alternative' routes showing the position of transition-states and intermediates. The axis can be in bond lengths or more usefully as Pauling bond orders.[50] Energy surfaces can be drawn with two of the axis representing changing bond lengths during the reaction and an implied third dimension of energy. These types of energy surface diagrams are referred to as More O'Ferrall-Jencks diagrams[43,44] (MOFJ) and have been used to rationalise observed chemical reactivity. In MOFJ diagrams the corners represent the extremes of the reaction where (a) are the starting materials, (b) are the products, (c) is the S_N1 intermediate (I) triple ion corner and (d) is the hypothetical, pentacovalent bonded species, Figure 4.5. Progress from reactants to products may occur in principle by a two-step mechanism involving the formation of an intermediate as in the classical S_N1 mechanism. Alternatively, the reaction may occur by a one-step process in which dissociation of the leaving group is concerted with the association of the nucleophile, the classical S_N2 mechanism.

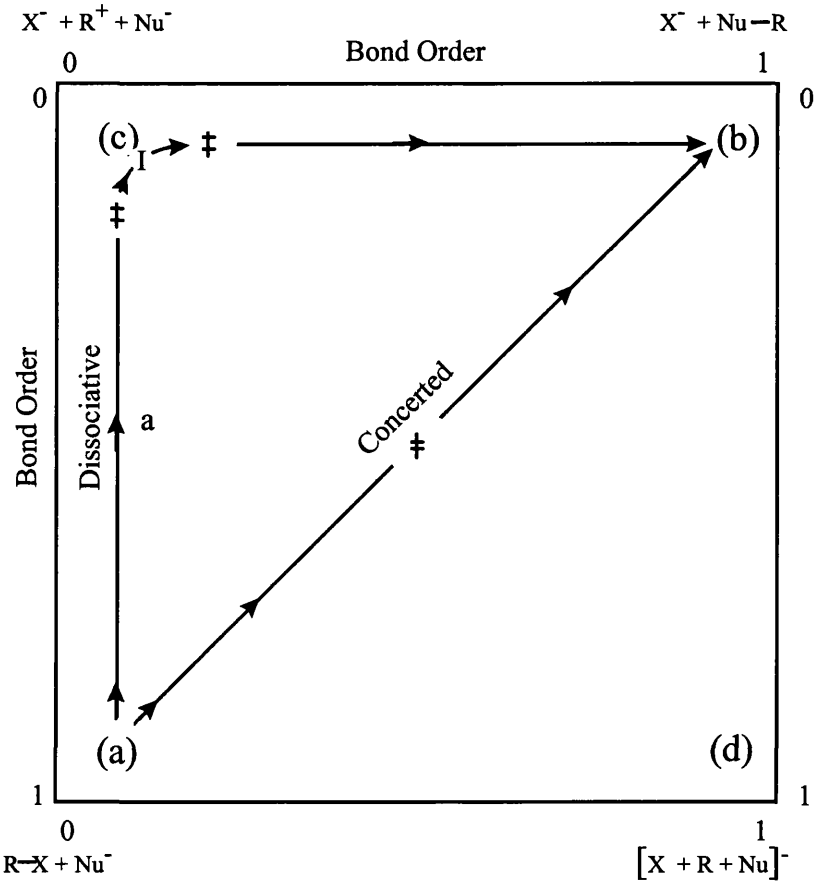


Figure 4.5: More O'Ferrall-Jencks energy surface diagram. The corners represent the extremes of the reaction where (a) are the starting materials, (b) are the products, (c) is the S_N1 intermediate (I) triple ion corner and (d) is the hypothetical, pentacovalent bonded species.

The MOFJ diagrams can be used to describe how the mechanism of nucleophilic substitution can change from S_N2 to S_N1 as the nature of the substrate varies. Due to the difficulty of studying borderline reactions experimentally, the exact details at this transitional point are still unclear; however, various possibilities have been put forward. The first is that both mechanisms can coexist within a borderline region (Hughes and Ingold)[34] with one being predominant over the other, while at the borderline both are present and equal. Experimentally, the suggested evidence comes from the concurrent stepwise reactions of *p*-methoxybenzyl derivatives, the solvolysis of acyl chlorides and the substitution reaction of (*p*-methoxybenzyl) dimethylsulfonium chloride.[45] Another possibility is that the reaction can jump discontinuously from one mechanism to the other, with the S_N1 intermediate being quite different from the S_N2 transition-state. The experimental evidence for the 'discontinuous jump' in mechanism comes from the reactions of various nucleophiles with 1-phenylethyl derivatives.[46] There also exists the possibility that one mechanism changes smoothly into the other with the S_N1 intermediate closely resembling the S_N2 transition-state (Doering and Zeiss).[47] Experimental evidence for this possibility is taken from the nucleophilic substitution by azide anion on ring substituted cumyl derivatives.[48]

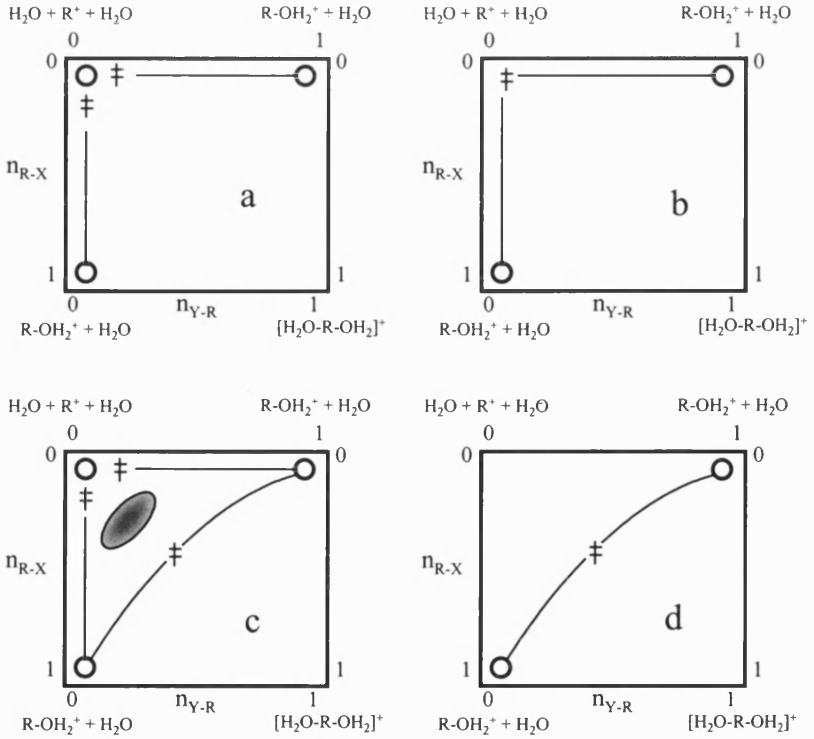


Figure 4.6: More O'Ferrall-Jencks diagrams depicting mechanistic change. The changeover from S_N2 (a) to S_N1 (d) reaction mechanism could occur by an S_N2 (b) reaction with essentially the same structure as the S_N1 intermediate or an S_N1 and S_N2 (c) pathway running concurrently.

By making various substitutional changes on the system it is possible to computationally study the changeover point in mechanism. The changeover from S_N2 (a) to S_N1 (d) reaction mechanism could occur by an S_N2 (b) reaction with essentially the same structure as the S_N1 intermediate or an S_N1 and S_N2 (c) pathway running concurrently, Figure 4.6. The Doering and Zeiss model can be represented by a gradual, continuous changeover from (b) to (a).

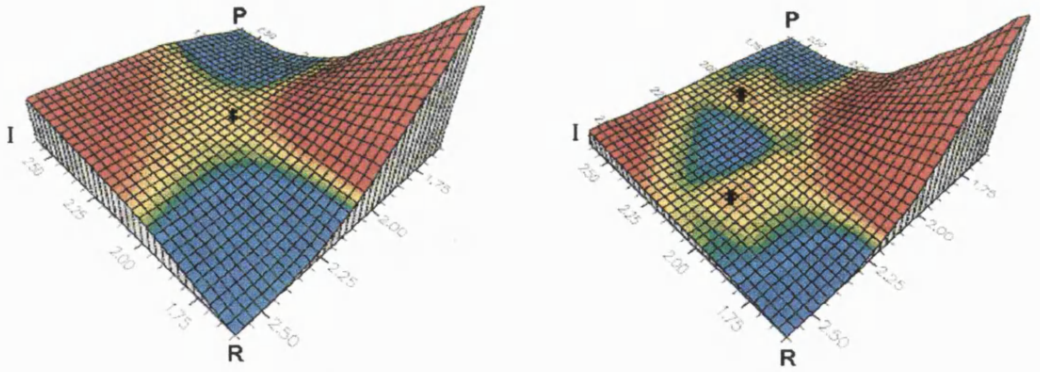
4.4 Previous work

A series of degenerate S_N2 reactions for a range of α-substituted alkyl protonated alcohols, Equation 4.5, where R_1 = H or Me and R_2 = H, Me or MeO were previously investigated by Barnes, Wilkie and Williams.[49] They calculated gas-phase barrier heights and energy surfaces using the AM1 semi-empirical MO method.

$$H_2O + CHR_1R_2OH_2^+ \rightarrow {}^+OH_2CHR_1R_2 + H_2O \qquad \text{Equation 4.5}$$

They concluded from their results that the stability of the carbocation intermediate increases as the number of electron-donating groups on the α-carbon is increased. This is reflected in the diminishing carbon-oxygen bond orders of the symmetrical transition-states.

AM1 *in vacuo* energy surfaces were calculated for the degenerate S_N2 reaction for (a) $MeCH_2OH_2^+$ and (b) $MeOCH_2OH_2^+$ as a function of the making and breaking C--O bond distances (Å). The surfaces appear to be similar in their overall form, but the coloured energy contours revealed a striking dissimilarity in the region of the S_N2 transition-state and S_N1 intermediate. The methoxymethyl substrate contained a shallow well (0.6 kJ mol^{-1}) at the same location of the ethyl transition-state. It appeared that the changeover in mechanism occurred in a structurally continuous way as described by the Doering and Zeiss model.



(a) $MeCH_2OH_2^+$ (b) $MeOCH_2OH_2^+$

Figure 4.7: AM1 *in vacuo,* energy surfaces for (a) $MeCH_2OH_2^+$ and (b) $MeOCH_2OH_2^+$. Low energies are coloured blue while high energies are coloured red. **R** represents the reactant corner, **P** the product corner, **I** the intermediate corner and ‡ marks the position of the transition-state.

The role of solvation was not included by the above theoretical study. The influence of solvent on organic reactions can be both complex and subtle. The ability to dissolve the reactants and bring them into intimate contact with one another is only one aspect of the role of the solvent in a reaction. Solvation involves the electrostatic interaction between solvent molecules and solvated species. The differential solvation of a transition-state can lower the activation energy for its formation and so increase the rate of reaction. The effectiveness of solvation may also depend to some extent on the need for stabilisation. The more internal stabilisation that occurs by inductive or resonance charge dispersal then the smaller contribution made by the solvent. The role of solvent is an important property to be included in any model that attempts to offer insight into experimental data which is mainly carried out in solution.

4.5 Methods

$$H_2O + CR_1R_2R_3OH_2^+ \rightarrow {}^+OH_2CR_1R_2R_3 + H_2O \qquad \text{Equation 4.6}$$

The bimolecular (S_N2) and unimolecular (S_N1) solvolysis reactions of increasingly alkylated protonated alcohols was studied, Equation 4.6, where R_1= H, Me or MeO, R_2 = H or Me, and R_3 = H or Me. The AM1 semi-empirical Hamiltonian utilising a continuum model, COSMO in MOPAC93, was used to model the reaction in water.[11,12] Each transition structure was characterised as having only one imaginary frequency. The kinetic isotope effects for the series are evaluated from AM1/COSMO geometries and Cartesian force constants using the CAMVIB and CAMISO programs.[28]

4.5.1 Pauling bond order

A useful way to interpret bond length displacements is to express them as Pauling bond orders n according to the relation $n = \exp[(r_1 - r_n)/c]$, where r_1 and r_n are lengths for bond order of 1 and n.[50] Although a value of 0.3 for the constant c has been widely employed, it has been noted elsewhere that this value correlates to changes in bond orders > 1, due largely to changes in the degree of π-bonding, that a different value is more appropriate for changes in bond order < 1 which are largely due to changes in the degree of σ-bonding.[51] The proportionality constant c was calculated from the transition structure for the degenerate reaction, $H_2O + CH_3OH_2^+ \rightarrow CH_3OH_2^+ + H_2O$, in which $n = 0.5$ by definition and where r_1 and r_n are the bond length in $CH_3OH_2^+$ and in the S_N2 transition-state, respectively. A value of 0.67 was obtained for c.

4.6 Results

There exists two distinct mechanisms for the displacement of H_2O from a saturated carbon centre, Equation 4.6, by a nucleophile, H_2O. The enthalpy of activation for the S_N2 mechanism was defined as:

$$\Delta H^{\ddagger}_{(S_N2)} = \Delta H_{f(S_N2\ TS)} - \{\Delta H_{f(H_2O)} + \Delta H_{f(ROH_2^+)}\} \text{ / kJ mol}^{-1} \qquad \text{Equation 4.7}$$

The enthalpy of activation for the S_N1 mechanism was defined as:

$$\Delta H^{\ddagger}_{(S_N1)} = \Delta H_{f(S_N1\ TS)} - \Delta H_{f(ROH_2^+)} \text{ / kJ mol}^{-1}. \qquad \text{Equation 4.8}$$

Table 4.1 contains the energies of activation for both mechanistic pathways for the degenerate reaction, Equation 4.7 and Equation 4.8, together with the transition-state

carbon-oxygen bond lengths. Table 4.2 contains the transition-state Pauling bond orders for both the S_N2 and S_N1 mechanistic pathways together with the 2° α-deuterium kinetic isotope effects, k_H/k_D, at 298.15K for a single deuterium atom.

Table 4.1: AM1/COSMO calculated energy barriers and bond lengths for the S_N2 and S_N1 mechanisms.

Reactant				$\Delta H^{\ddagger}_{(S_N2)}$ / kJ mol^{-1}	$\Delta H^{\ddagger}_{(S_N1)}$ / kJ mol^{-1}	S_N2 C - O/Å	S_N1 C-- O/Å
R_1	R_2	R_3					
H	H	H	Me	91.7	136.1	1.94	3.30
Me	H	H	Et	84.3	104.6	2.08	2.90
Me	Me	H	^{i}Pr	57.6	51.1	2.30	2.73
Me	Me	Me	^{t}Bu	31.5	18.1	2.49	2.36
MeO	H	H	MeO	51.2	54.2	2.11	3.01
MeO	Me	H	MeOMe	31.2	24.3	2.24	2.18

Table 4.2: AM1/COSMO calculated Pauling bond orders and kinetic isotope effects for the S_N2 and S_N1 mechanisms.

Reactant	S_N2 n_{C-O}	S_N1 n_{C-O}	S_N2 k_H/k_D	S_N1 k_H/k_D
Me	0.50	0.07	1.01	1.36
Et	0.41	0.12	1.10	1.28
^{i}Pr	0.29	0.15	1.17	1.25
^{t}Bu	0.22	0.27	*	*
MeO	0.39	0.10	1.16	1.29
MeOMe	0.32	0.35	1.15	1.20

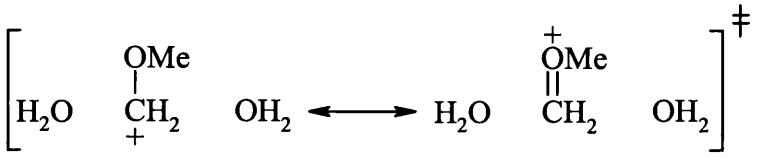
4.7 Discussion

4.7.1 Barrier heights

As the substitution of hydrogens by electron-donating methyl groups at the α-carbon increases, stabilisation of the carbocation occurs. This leads to a decrease in both the S_N1 and S_N2 activation barrier heights, Table 4.1. The effect of substituting a single methyl group for a proton at the α-carbon, to give a primary carbocation, on the S_N2 barrier height is to lower it by only 7.4 kJ mol^{-1} as compared to 31.5 kJ mol^{-1} for the S_N1

barrier height. The further addition of a methyl group for another proton, to give a secondary carbocation, reduces the barrier heights further, by 26.7 and 53.5 kJ mol^{-1} for the S_N2 and S_N1 barrier heights, respectively. The further addition of a methyl group for the remaining proton to give a tertiary carbocation, lowers the S_N2 and S_N1 barrier heights by approximately the same amount, 26.1 and 33.0 kJ mol^{-1}, respectively. The overall effect of increased alkylation is to decrease the activation energy due to the increase in carbocation stabilisation it offers.

The effect of increasing alkylation is far greater for the S_N1 than for the S_N2 mechanism due to the relative needs for transition-state stabilisation. The carbocation-like fragment of the S_N2 transition-state is sandwiched between two water molecules (nucleophile and leaving group) that stabilise it and so the addition of the methyl-group substituents has less effect on the α-carbon. The S_N1 transition-state carbocations are only stabilised by one water molecule (the leaving group) and so the addition of methyl groups at the α-carbon has a relatively larger effect. The substitution of a methoxy group for a proton dramatically reduces the barrier heights for the S_N2 and S_N1 barriers by 40.5 and 81.9 kJ mol^{-1}, respectively. The addition of the α-methoxy group causes an increase in reactivity,[52,53] since its net effect is to donate electrons to the developing positive charge at the transition-state by contributions to structures such as



Equation 4.9

The changeover in mechanism from S_N2 to S_N1 for the alkyl series occurs between protonated ethanol, S_N2, and protonated *iso*-propanol, S_N1. The addition of the alkyl substituents cause a relative decrease in the S_N2 reactivity at the α-carbon.[54,55] This shows the same changeover point as Ingold in his studies with alkyl bromides. The status of the *t*-butyl group is consistent with it being inert towards the S_N2-reaction: steric repulsion between α-carbon substituents and the incoming nucleophile requires that in order for the reaction to occur the C-O bonds must become greatly elongated.[56]

4.7.2 Kinetic isotope effect (KIE)

The magnitude of a secondary α-deuterium kinetic isotope effect for nucleophilic substitution depends upon the nature of the nucleophile, leaving group, α-substituent and solvent, but is generally used to discriminate between S_N2 and S_N1 mechanisms for a reaction; the former typically shows values close to unity (either normal or inverse) while the later usually show larger normal values.[57] The KIE for the substitution of a single deuterium for a single proton, Equation 4.10, where D is the deuterated analogue at the α-C were calculated at 298.15K, Table 4.2.

$$H_2O + CDR_1R_2OH_2^+ \rightarrow {}^+OH_2CDR_1R_2 + H_2O \qquad \text{Equation 4.10}$$

As the stability of the S_N2 transition-state increases and it becomes more carbocation-like in character, so the transition-state bond orders, $n_{C\text{-}O}$, decrease. This is indicative of a progressively 'looser' S_N2 transition-state. Examination of the calculated α-deuterium isotope effects for the S_N2 mechanism shows a value that it is close to unity for protonated methanol and then increases as the series is descended. This is due to the increasing stability of the carbocation which leads to a 'looser' S_N2 transition-state. Examination of the S_N1 bond orders shows the reverse trend. As the stability of the carbocation increases the S_N1 transition-state becomes tighter and the KIE decreases in value as the series is descended. The S_N2 transition-states for the methoxy alkyl substituents give KIE values with magnitudes similar to those usually associated with for S_N1 mechanisms. These large normal values are due to an open, exploded transition-state with a large amount of carbocation character which has freedom for the bending motion for the C-H bonds of the formaldehyde moiety.[58]

The KIE values for the S_N2 transition-state are small for a 'tight' transition-state such as in the methyl substrate or relatively large for a 'loose' transition-state such as in the *iso*-propyl substrate. A value in this range is usually attributed to a S_N1 and not S_N2 mechanism. This implies that when applying KIE to the elucidation of mechanism, great care must be taken. The theoretically calculated results are able to reveal results that are not accessible via experimental techniques. The kinetic isotope effects for reactions, such as the S_N1 reaction of a methyl or ethyl compound, and the S_N2 reaction for the *iso*-propyl or *t*-butyl compound can be calculated but are unobserved experimentally.

4.7.3 Energy surfaces

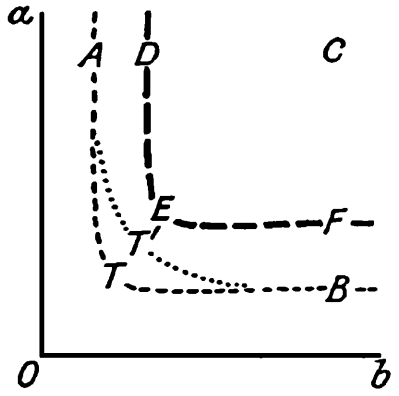


Figure 4.8: Ingold's imaginary potential surface showing the possible pathways of mechanistic change.

An early attempt to explain nucleophilic substitution was proposed by Ingold *et al.* with the aid of an imaginary potential surface, Figure 4.8.[34] They assumed the surface to have two valleys, A and B, usual for a gas phase reaction with a transition-state at T, but also a third valley at C is present (due to ion solvation) together with the pair of ridges DEF. Increasing the electron donation to the reaction zone, they expected an increasing gradient, negative in the directions 0*a* and 0*b*, to be imposed on the original surface. This would have the effect of raising the energy of activation AT, forcing the reaction path ATB to undergo only a slight displacement away from the origin surface. At a later stage along the series, when the increasing tilt has depressed the ridges DEF and largely flattened the hill TE, the reaction path AT'B will spread out much more widely. This would lead to a reduced increase and possibly, an actual decrease in activation energy. In certain cases the decrease may set in before the reaction path has become so far turned that the representative point runs into valley C, a unimolecular reaction. Ingold's imaginary surface could be regarded as an early 'simple' version of a potential energy surface used to describe possible mechanisms.

The energy surface proposed by Ingold were imaginary, however, by using modern computational methods it is possible to calculate such surfaces for real reactions in solution. The energy surface is calculated as a function of the carbon-nucleophile and carbon-leaving group bond length or bond order distances, Figure 4.9. As the reactions studied are degenerate, only half of the energy surface is calculated and the other half is taken as its 'reflection'. The co-ordinates of the energy surface are plotted using the UNIRAS package and the points are interpolated to give a uniform surface. The colours

represent different energy levels; low energies are coloured blue while high energies are coloured red.

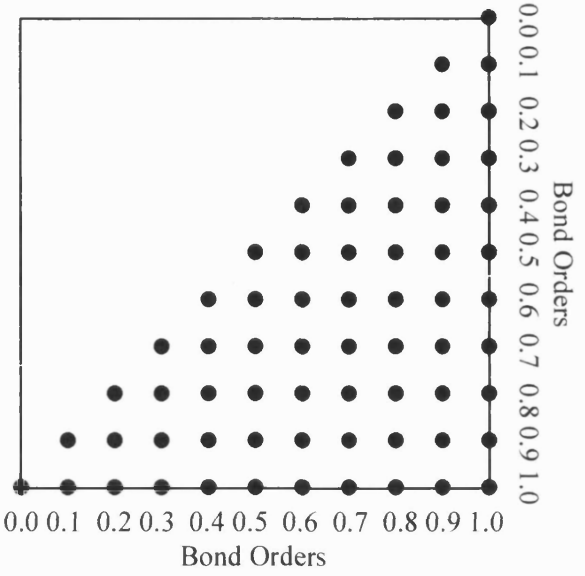


Figure 4.9: The energy surface template for a degenerate reaction. The co-ordinates are formed by pairs of values for the bond making and breaking orders.

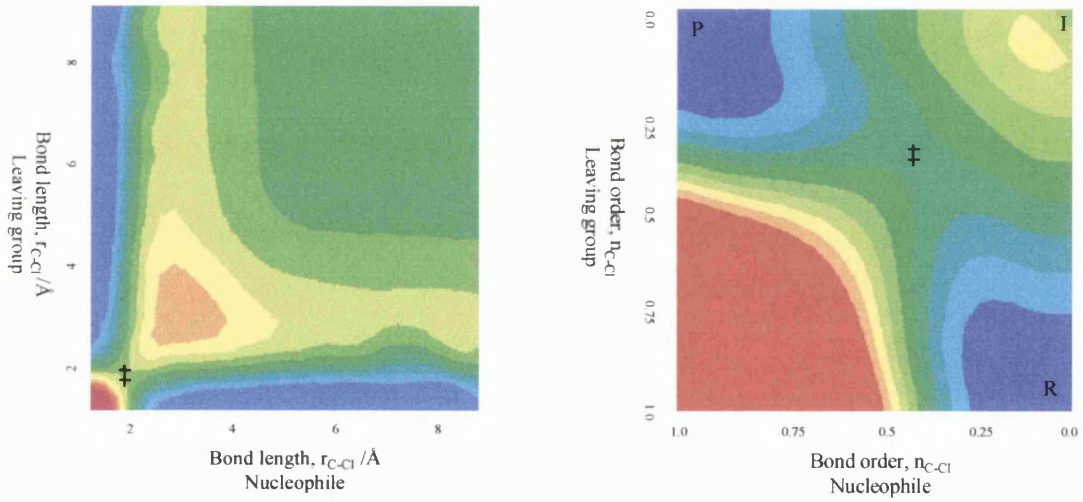


Figure 4.10: AM1/COSMO energy surface of ethylchloride/chloride anion *in aquo* shown in bond lengths and bond orders. Low energies are coloured blue while high energies are coloured red. ‡ represents the position of the transition-state, **I** represents the intermediate corner while **R** and **P** the reactants and products, respectively.

The imaginary surface proposed in the 1930's can be compared to a calculated AM1/COSMO energy surface for the degenerate ethylchloride/chloride anion *in aquo*, Figure 4.10. The similarities of the surfaces, the predicted ridges, should be noted. The conversion of a surface from bond lengths to bond orders dramatically emphasises the

usefulness of such energy surfaces. In the bond energy surface, the points of interest (transition-states and intermediates) are all concentrated in the bottom left corner whereas in the bond order surface points of interest are more predominant and spread across the surface. The bond order surface shows clearly that there is a possibility of two reaction paths (S_N1 and S_N2) while the bond length surface shows only one clearly even though the same 'raw' data is used. Visualisation of the third, energy dimension is possible with the reactants at the front, the products at the back and the left hand side corner representing the intermediate corner. This is the standard orientation of energy surfaces throughout this thesis.

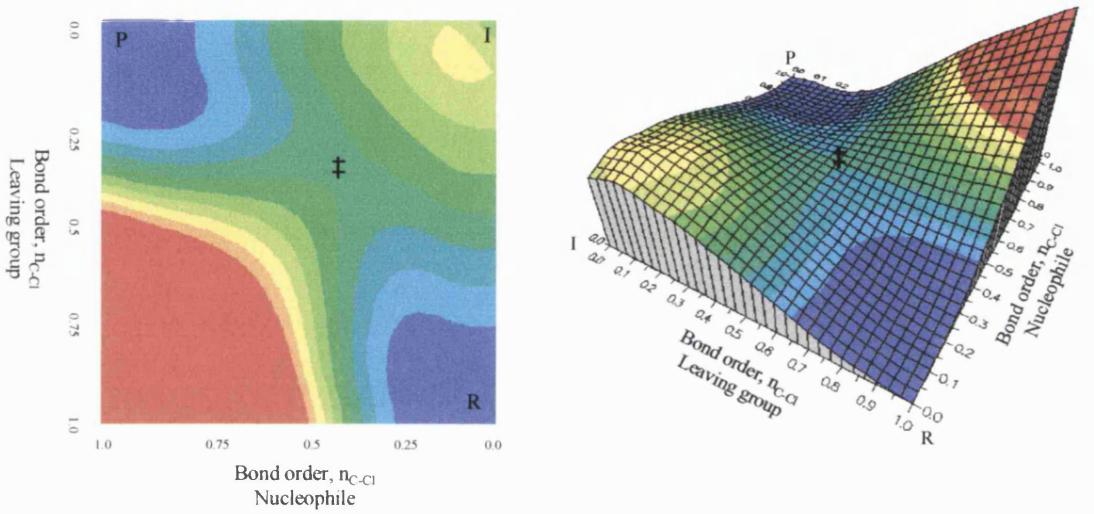


Figure 4.11: AM1/COSMO energy surface of ethylchloride/chloride anion *in aquo* shown in 2 and 3D. Low energies are coloured blue while high energies are coloured red. ‡ represents the position of the transition-state, **I** represents the intermediate corner while **R** and **P** the reactants and products, respectively.

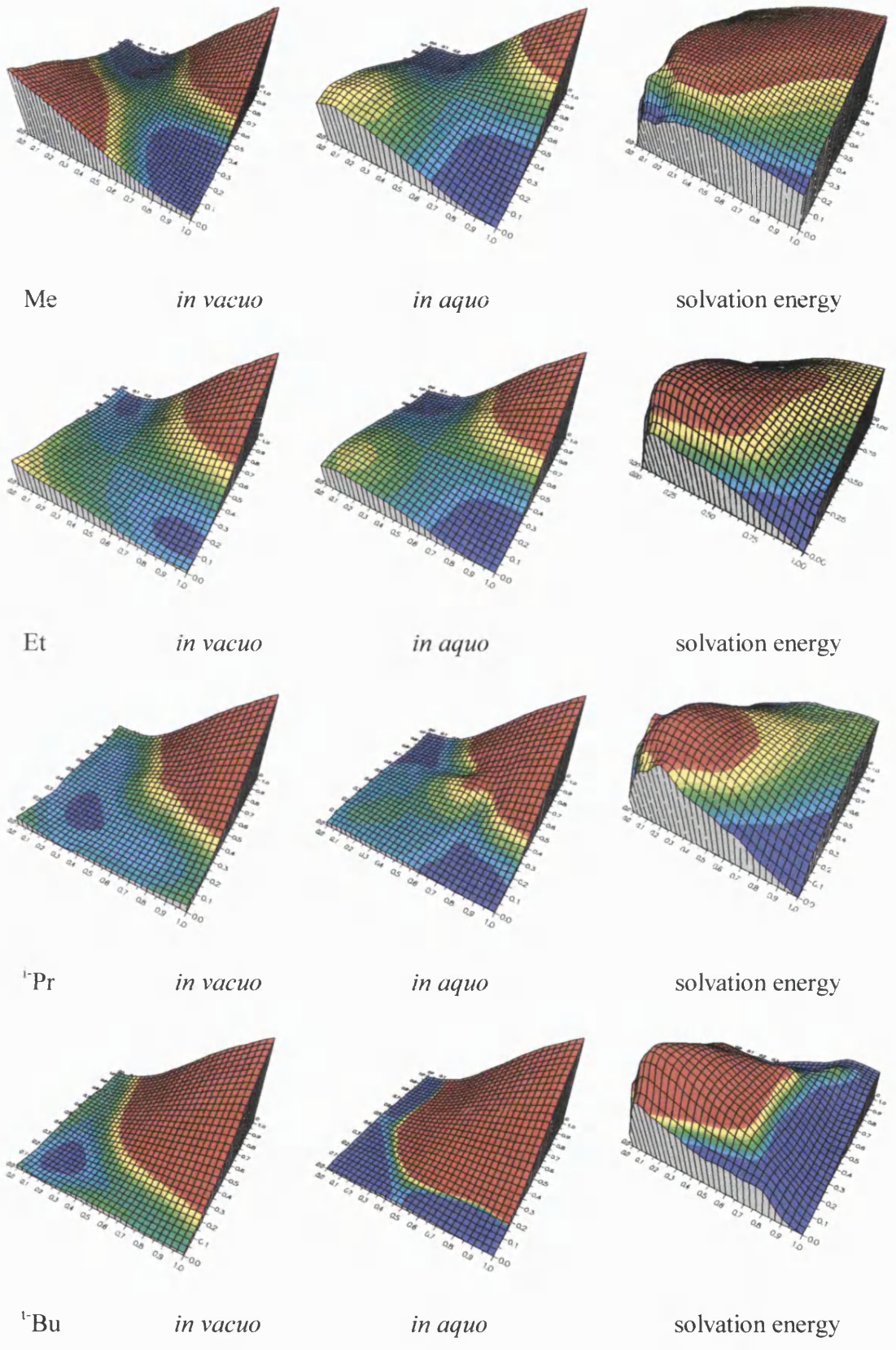
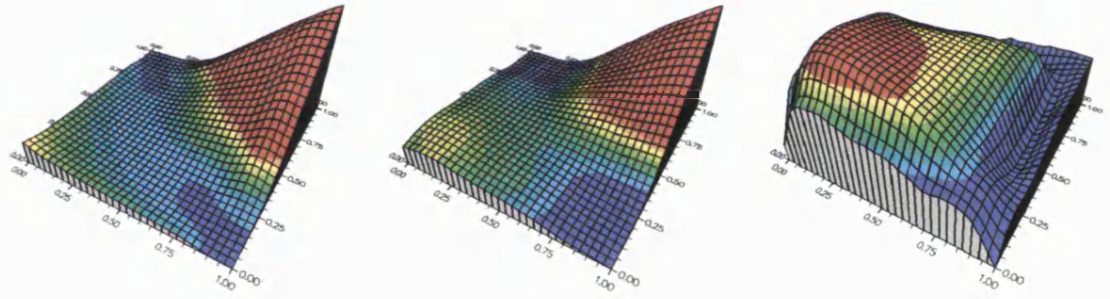
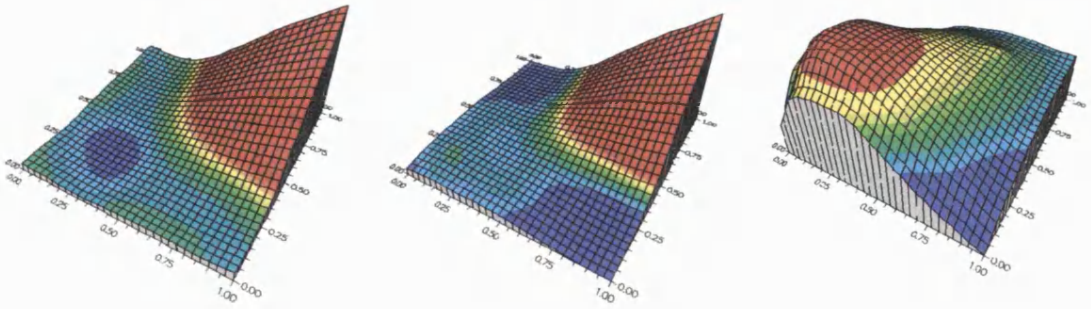


Figure 4.12: *In vacuo, in aquo* and solvation energy surfaces. Low energies are coloured blue while high energies are coloured red.

The energy surfaces for the degenerate reactions are calculated for the solvolysis reactions of increasingly alkylated protonated alcohols, Equation 4.6. For comparison, the *in vacuo, in aquo* and solvation energy surfaces are calculated. Whether the nucleophile is neutral, water, or anionic, chlorine anion, the energy surfaces produced are very similar. As the series, methyl, ethyl, *iso*-propyl *t*-butyl is descended, the S_N1 corner of the *in vacuo* energy surfaces becomes more stable as the inductive effect of the substituent increase and the S_N2 transition-state slides towards the S_N1 corner, Figure 4.12. As the same series is descended for the *in aquo* energy surfaces, the S_N1 corner becomes dramatically stabilised and eventually plateaus out. From the solvation energy surfaces it can be seen that the reactants and products are stabilised greatest and then to a lesser extent the S_N1 corner. This is due to both the reactants and products containing localised charges that are solvated to a greater extent than if the charge is spread. The exception to this is the protonated methanol solvated cation solvation energy surface whose S_N1 corner (the methyl cation) is stabilised to the greatest extent. This may be due to the small cation interacting strongly with the solvent or that COSMO has overestimated the solvation stabilisation energy of the methyl cation.



MeO *in vacuo* *in aquo* solvation energy



MeOMe *in vacuo* *in aquo* solvation energy

Figure 4.13: *In vacuo, in aquo* and solvation energy surfaces. Low energies are coloured blue while high energies are coloured red.

The effect on the energy surface of adding a methoxy group to the *in vacuo* surface is to lower the energy of the S_N1 intermediate corner, Figure 4.13. The further addition of a methyl group flattens the surface even more. The effect of solvent is to stabilise both the S_N1 intermediate and pentacovalent corners. The *in vacuo* energy surface reported by Barnes, Wilkie and Williams[49] for the methoxymethyl substrate appeared to contained a shallow well (0.6 kJ mol^{-1}) at the same location as the ethyl transition-state, implying that the changeover in mechanism occurred by a structurally continuous way as described by the Doering and Zeiss model. The present calculations show no shallow well for either the ethyl or MeO energy surfaces, but only a S_N2 saddle point. The shallow well was probably an artefact of the method used in the previous work.

The solvated energy surfaces all show two distinct mechanistic pathways as described by Ingold and Hughes, one of which is in ascendance over the other. The methyl and ethyl substrates are not sufficiently stabilising to allow the S_N1 route to be viable and so the S_N2 route is favoured. The *iso*-propyl and *t*-butyl substrates are sufficiently powerful electron-donors for the stabilisation of the S_N1 pathway to be lower in energy than the S_N2 pathway. The addition of methyl groups to the reacting α-carbon is to great a perturbation of the energy surface to show a balanced situation between two coexisting, concurrent mechanisms. The Winstein, Doering and Zeiss, Bentley and Schleyer and Sneen descriptions of mechanisms either see the S_N1 or S_N2 as extremes of a single merging mechanism or a discontinuous jump from one mechanism to the other. These are not seen in any of the energy surfaces.

4.8 Conclusion

The study of the protonated alcohols gave the same trend as Ingold *et al.* solvolysis study of a series of alkylated bromides showing the same point of mechanistic change (between the ethyl and *iso*-propyl substrates). The KIE were in line with predictions; small KIE for tight transition-state and large KIE for loose transition-states. The KIE values for the S_N2 transition-state are small for a tight transition-state (methyl) or relatively large for a loose transition-state (*iso*-propyl) which would usually be attributed to the S_N1 mechanism. A 'tight' transition-state, regardless of whether the

reaction is S_N2 or S_N1, will give an inverse or small KIE whereas a 'loose' transition-state will give a 'large' normal KIE.

The AM1/COSMO continuum solvation model would appear to show that the solvent stabilises the reactants and products to the greatest extent followed by the S_N1 intermediate corner. What can be seen throughout the series is that two distinct, competing mechanistic pathway exist; Ingold's S_N2 and S_N1 mechanistic pathways. The alternative mechanistic descriptions for changeover such as the Doering and Zeiss' 'merging' mechanistic pathway or a discontinuous jump from one mechanism to the other are not seen in any of the energy surfaces. The conclusion drawn by Barnes, Wilkie and Williams that the transition changeover in mechanism occurs by Doering and Zeiss' structural hypothesis appears to be incorrect. Their surfaces did not encompass enough of the total energy surface to show alternative mechanistic possibilities.

Solvent effects are known to have profound influence on the rate of reactions. The bulk and specific interactions between solvent molecules and the dissolved ions is an important property that requires modelling. COSMO is a continuum solvation model that attempts to account for the bulk water interactions but not specific molecular interactions. This means that it can not take account of any specific solvent-solute interactions such as hydrogen bonds and the stabilisation that it offers. This may lead to an under- or over-estimation of any solvent stabilisation. Experimental and theoretical evidence has suggested that primary carbocations are transformed by hydration into alkoxonium ions solvated by four or five additional water molecules. The *t*-butyl cations are more weakly solvated, however, they still bind four water molecules in their hydration sphere.[59] The inclusion of specific solvent-solute interactions by the utilisation of a hybrid QM/MM method would give insight into the specific stabilisation effect of the solvent. A semi-empirical or *ab initio* QM/MM method such as the GRACE/CHARMM suite of programs should be used to study an appropriate system and this is detailed in Chapter 9.[60,15]

Chapter Five

Theoretical Modelling of S_N2/S_N1 Reactivity of Benzyl Derivatives: The Identity Reaction of Substituted Benzyl Chlorides

Chapter 5 - Theoretical Modelling of S_N2/S_N1 Reactivity of Benzyl Derivatives: The Identity Reaction of Substituted Benzyl Chlorides

5.1 Introduction

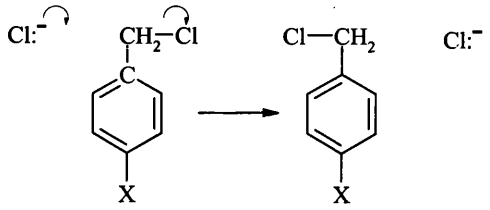
Benzyl derivatives are a highly versatile group of synthetic reagents that are commonly used to mask alcohols and carboxyl functionality.[61] Nucleophilic substitution reactions of benzyl derivatives have been studied for most of this century and it has been noted that the mechanism depends upon the nature of the substituent on the aromatic ring.[62] As a consequence, the study of these reactions has been central to the development of the mechanistic framework for nucleophilic substitution reactions at aliphatic carbon.

A long-standing goal of chemists is to develop models that explain the effects of changing structure on the rates and mechanisms of organic reactions. It is especially challenging to explain the results of studies of solvolysis reactions and nucleophilic substitution reactions of benzyl derivatives.[63] In the past this data has been extensively analysed and interpreted.[64] There is general agreement that there are both changes in mechanism, from stepwise (S_N1) to concerted (S_N2), and transition-state structure as the substituents on the aromatic ring are made electron-withdrawing relative to *p*-MeO.

In Chapter 4, the mechanistic changeover of identity reactions of protonated alcohols with water in aqueous solution was investigated. It was proposed that the reaction always occurred by two competing pathways but that one was predominant. Mechanistic changeover would occur at a borderline where the barriers to both mechanisms were approximately equal and both mechanisms would therefore coexist with neither being in dominance. However, the effect of the addition of a methyl group to the reaction was too great a perturbation on the energy surface to allow this borderline to be studied in a detailed way.

5.2 Methods

In order to gain a better understanding of the intrinsic characteristics of nucleophilic substitution reactions at carbon centres, and in particular the effect of various substituents, we studied the nucleophilic substitution 'identity' reactions of substituted benzyl chlorides, Figure 5.1. This is an ideal system to investigate due to the benzene ring acting as a structural probe on which substituents can be placed. These substituents can then have an electronic effect on the reacting carbon centre which is far removed from that centre therefore causing only a slight perturbation on the energy surface.



X = NMe_2, NMeH, NH_2, OH, MeO, Me, $CHCH_2$, H, Br, CHO, CN, and NO_2.

Figure 5.1: Substituted benzyl chloride 'identity' reaction where X is the *p*-substituent varied.

It is essential to adopt a calculational method capable of providing realistic descriptions of the relative energetics of families of reactions involving substituted substrates even if not the absolute barriers. The AM1 semi-empirical Hamiltonian utilising COSMO,[11] a continuum model in MOPAC93,[12] is used to mimic solvation. The S_N2 and S_N1 activation energies are calculated for each mechanism. The lowest energy orientation for the transition states is obtained with the C–Cl bond vector perpendicular to the benzene ring. Such an orientation in the transition-state allows for the overlap of the chlorine p-orbitals with the π-system in the benzene ring lowering the energy of the system. Each transition-state structure is characterised as having only one imaginary frequency. The CAMVIB and CAMISO programs is used to compute the kinetic isotope effects at 298.15K.[28]

5.3 Results

The activation energy for the nucleophilic substitution identity reactions of *p*-substituted benzyl chlorides for the S_N2 and S_N1 mechanisms are calculated for a series of twelve different substituents, Table 5.1. The enthalpy of activation for the S_N2 mechanism is defined as

$$\Delta H^{\ddagger}_{(S_N2)} = \Delta H_{f(S_N2\ TS)} - \{\Delta H_{f(XBzCl)} + \Delta H_{f(Cl^-)}\} \text{ / kJ mol}^{-1}. \qquad \text{Equation 5.1}$$

The enthalpy of activation for the S_N1 mechanism is defined as

$$\Delta H^{\ddagger}_{(S_N1)} = \Delta H_{f(S_N1\ TS)} - \Delta H_{f(XBzCl)} \text{ / kJ mol}^{-1}. \qquad \text{Equation 5.2}$$

The transition-state bond lengths (Å) for the carbon-chlorine, C-Cl, distance for the various mechanisms are given together with the Pauling bond orders which are calculated in the same manner as described in Chapter 4.

Table 5.1: AM1/COSMO calculated activation energies, kJ mol^{-1}, bond lengths and Pauling bond orders for the S_N2 and S_N1 mechanisms.

Substituent X	$\Delta H^{\ddagger}_{(S_N2)}$	$\Delta H^{\ddagger}_{(S_N1)}$	S_N2 C-Cl/Å	S_N1 C-Cl/Å	S_N2 $n_{C\text{-}Cl}$	S_N1 $n_{C\text{-}Cl}$
NMe_2	118.6	87.7	2.32	2.69	0.45	0.26
NMeH	117.6	83.3	2.32	2.70	0.45	0.25
NH_2	120.6	77.2	2.32	2.75	0.45	0.20
OH	121.0	111.3	2.30	2.91	0.46	0.19
MeO	121.0	114.4	2.29	2.95	0.46	0.17
Me	121.8	125.3	2.28	3.01	0.47	0.16
$CHCH_2$	122.3	126.6	2.28	3.03	0.47	0.16
H	123.0	130.6	2.28	3.10	0.47	0.14
Br	125.0	146.7	2.27	3.26	0.48	0.11
CHO	125.5	152.3	2.26	3.33	0.49	0.10
CN	126.1	151.6	2.26	3.35	0.49	0.09
NO_2	128.3	165.3	2.25	3.37	0.50	0.09

Table 5.2 contains the α-deuterium kinetic isotope effect, k_H/k_D, per deuterium atom at 298.15K together with the calculated transition state-reactant charge difference, Δq_{benzyl}, on the benzyl system. This is defined as the total charge on the benzyl group plus substituent at the transition-state minus the total charge on the benzyl group plus substituent of the reactants. The relative energy barrier populations are calculated using Equation 5.3, where $\delta(\Delta H)$ is the enthalpy difference between the two transition states, R is the ideal gas constant and T is the temperature.

$$\text{Relative population, } P = \exp\left\{\frac{-\delta(\Delta H)}{RT}\right\} \qquad \text{Equation 5.3}$$

Using this approach, the relative population of the lowest energy transition states is equal to one. The populations are then normalised, so that the sum of the populations $P(S_N2)$ and $P(S_N1)$ is equal to one. From these populations, the weighted energy barrier can be calculated using Equation 5.4.

$$\Delta H^{\ddagger}_{Wt} = P(S_N2) \times \Delta H^{\ddagger}_{(S_N2)} + P(S_N1) \times \Delta H^{\ddagger}_{(S_N1)} \qquad \text{Equation 5.4}$$

The same procedure is used to calculate a weighted kinetic isotope effect. Table 5.3 contains the α-carbon-benzene ring carbon bond lengths, C_α-C_{AR}, for the reactants and at the transition-state for the various mechanisms together with Pauling bond orders. The proportionality constant $c = 0.3$ was used to calculate the Pauling π-double bond orders.

Table 5.2: AM1/COSMO calculated secondary α-deuterium kinetic isotope effects, k_H/k_D, benzyl group charges, Δq_{benzyl}, for the S_N2 and S_N1 mechanisms, weighted barrier heights, kJ mol^{-1}, and kinetic isotope effects.

Substituent X	S_N2 k_H/k_D	S_N1 k_H/k_D	S_N2 Δq_{benzyl}	S_N1 Δq_{benzyl}	$\Delta H^{\ddagger}_{Wt}$	weighted k_H/k_D
NMe_2	1.02	1.10	0.45	0.83	87.7	1.10
NMeH	1.02	1.11	0.46	0.83	83.3	1.11
NH_2	1.01	1.12	0.46	0.82	77.1	1.12
OH	1.02	1.15	0.45	0.83	111.5	1.15
MeO	1.03	1.16	0.43	0.84	114.8	1.15
Me	1.01	1.17	0.43	0.83	123.0	1.05
$CHCH_2$	1.02	1.17	0.42	0.83	122.4	1.04
H	1.02	1.19	0.42	0.83	123.3	1.03
Br	1.02	1.19	0.40	0.83	125.0	1.02
CHO	1.02	1.19	0.39	0.84	125.5	1.02
CN	1.02	1.20	0.38	0.84	126.1	1.02
NO_2	1.02	1.22	0.38	0.85	128.3	1.02

Table 5.3: AM1/COSMO calculated bond lengths and bond orders for the α-carbon-benzene ring carbon bond, C_α-C_{AR}, for the S_N2 and S_N1 mechanisms.

Substituent X	Reactant C_α-C_{AR}/Å	S_N2 C_α-C_{AR}/Å	S_N1 C_α-C_{AR}/Å	S_N2 $n_{C_\alpha\text{-}C_{AR}}$	S_N1 $n_{C_\alpha\text{-}C_{AR}}$
NMe_2	1.478	1.441	1.368	1.13	1.44
NMeH	1.478	1.441	1.367	1.13	1.44
NH_2	1.478	1.441	1.360	1.13	1.48
OH	1.478	1.445	1.370	1.11	1.43
MeO	1.479	1.449	1.378	1.11	1.40
Me	1.481	1.453	1.385	1.10	1.38
$CHCH_2$	1.480	1.454	1.384	1.09	1.38
H	1.481	1.454	1.382	1.09	1.39
Br	1.482	1.458	1.387	1.08	1.37
CHO	1.482	1.461	1.391	1.08	1.35
CN	1.482	1.460	1.388	1.08	1.37
NO_2	1.485	1.464	1.396	1.07	1.35

5.4 Discussion

5.4.1 Barrier heights

The calculated activation energies and transition-state bond lengths for the S_N2 mechanism show only small variations in energy in going from electron-donating to withdrawing substituents, Table 5.1. Substituted benzyl chlorides that have electron-donating substituents at the *para* position have a lower activation energy than those with electron-withdrawing substituents due to their ability to stabilise the developing positive charge. The activation energies for the S_N1 mechanism show that for the more powerful electron-donating substituents, a reduction in the barrier occurs as the carbocation transition-state is stabilised. As the series is descended, the more powerful electron-withdrawing substituents ceases to have this ability and so the transition-state becomes less stable. The decrease in carbocation stability lengthens the S_N1 bond at the transition-state. The weighted activation energies show a smooth, gradual increase as the *para* substituent is made more electron-withdrawing. Table 5.2 shows that the calculated charge difference, Δq_{benzyl}, of the benzyl group charge for the S_N2 decreases as the transition-state becomes tighter and therefore less carbocation-like.

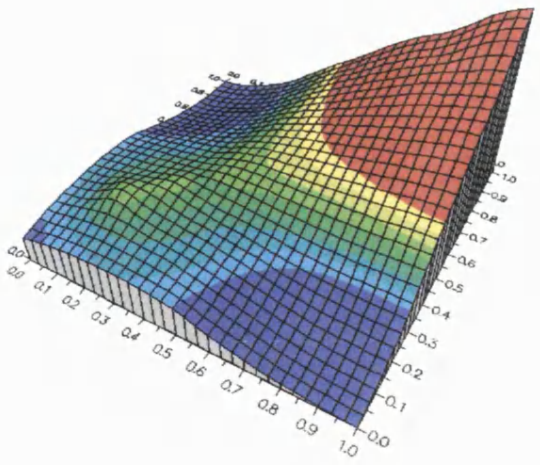
5.4.2 Kinetic isotope effect (KIE)

The experimental α-deuterium kinetic isotope effects for a series of benzyl chlorides with cyanide ion as the nucleophile varied as different substituent were varied. The cyanide and solvolysis reactions of *m*-chlorobenzyl chloride in 55% aqueous methyl cellosolve gave typical S_N2 kinetic isotope effects, close to unity, whereas the reactions of *p*-methoxybenzyl chloride had high values of 1.25-1.31. This was taken as evidence for a pathway involving a carbocation or an ion-pair mechanism. It was then inferred that kinetic isotope effect values between unity and that for the carbocation or ion-pair mechanism was due to the fact that a dual mechanistic pathway was in operation.[65] The effect of substituents on KIEs is to decrease their value as the substituents is made more electron-withdrawing.[66] For the bimolecular reaction of thiophenoxide nucleophile and substituted benzyl chlorides gave a k_H/k_D values for the *p*-methoxy, unsubstituted and *p*-nitro substituent of 1.061, 1.045 and 1.19, respectively.

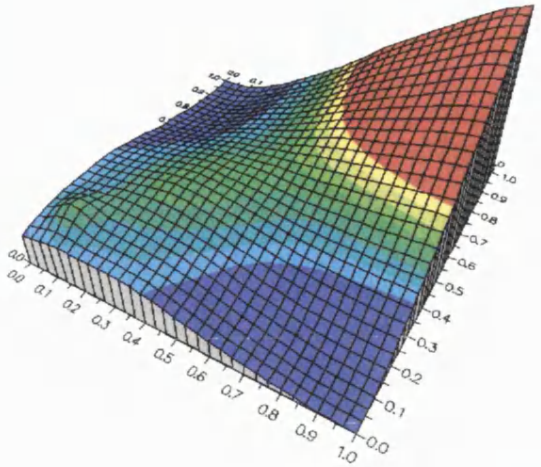
The KIE for the substitution of a deuterium for a proton at the α-C are calculated at 298.15K, Table 5.2. As the series is descended, the α-deuterium isotope effects for the S_N2 mechanism show an approximately constant KIE due to the lack of any substantial substituent effect in the S_N2 transition state as would be predicted for a constant transition state. If the effect of the substituent is substantial, then the variation in isotope effect would be greater.[66] The calculated values for the S_N2 mechanism are in the same region as the experimental S_N2 isotope effects for the reaction of various nucleophiles with substituted benzyl bromides[67] and the identity reaction of benzyl chloride.[68] As the series is ascended, the α-deuterium isotope effects for the S_N1 mechanism show a decreasing trend towards unity. This is what would be predicted by a varying transition-state which becomes tighter and more S_N2-like as the electron-donating ability of the substituent is increased.

5.4.3 The S_N1/S_N2 mechanistic borderline

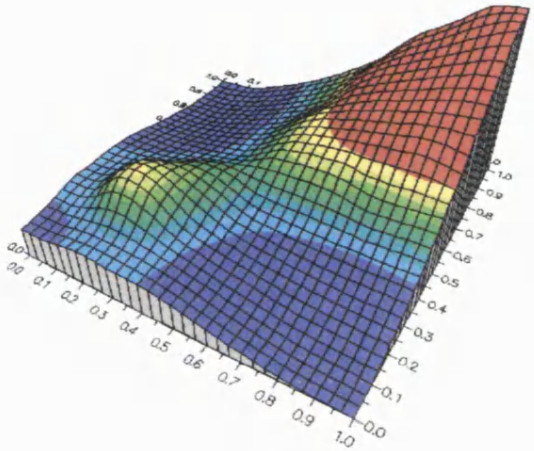
Energy surfaces can be drawn with two of the axes representing changing bond lengths during the reaction and an implied third dimension of energy. The axes are in Pauling bond orders and are commonly referred to as More O'Ferrall-Jencks[43,44] diagrams. The mechanism of nucleophilic substitution can change from S_N1 to S_N2 as the nature of the substrate varies. In this study it was desirable to investigate the energy surfaces of the twelve substituted benzyl chlorides. The range of substituents chosen act as a sensitive probe to investigate the changeover point in mechanism as their introduction causes only a slight, gradual perturbation of the potential energy surface and so the changeover point can be investigated. The AM1/COSMO energy surfaces are generated for the nucleophilic displacement reaction of all twelve substituents as described in Chapter 4. The energy surfaces show a maximum in the pentavalent corner. The S_N2 transition-state can be seen on the surfaces going straight across the energy surface from reactants to products with a saddle point at the middle of the surface. The saddle point shifts away from the associative corner to the dissociative corner as the substituent is made more electron-donating. Near the dissociative corner, a slight hilltop can be seen in the vicinity of the S_N1 transition-state and intermediate. The S_N1 transition-state shifts away from the dissociative corner back towards the reactant corner as the substituent is made more electron-donating.



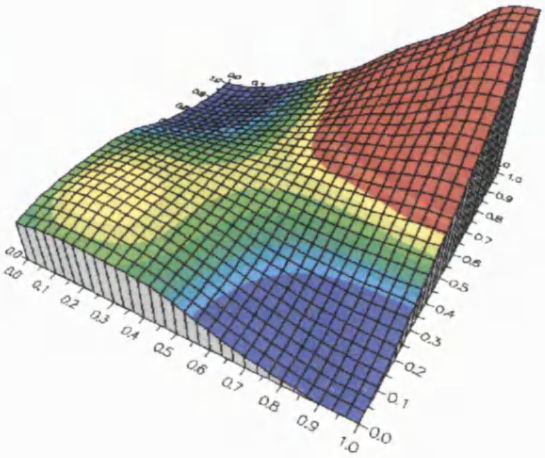
NMe_2



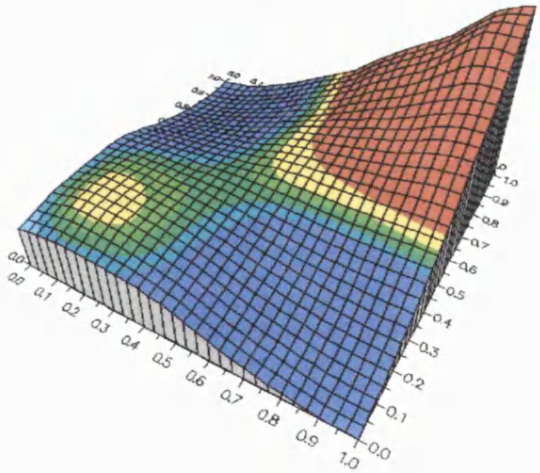
NMeH



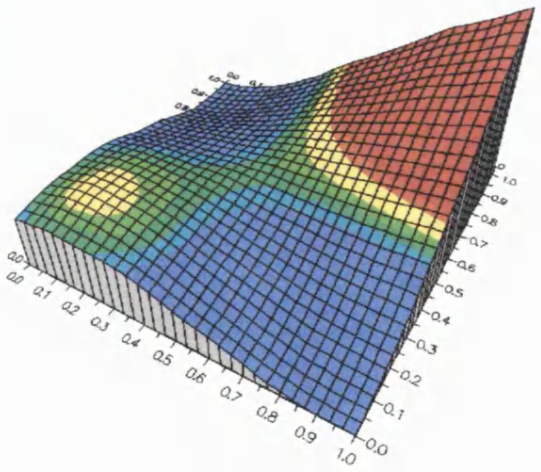
NH_2



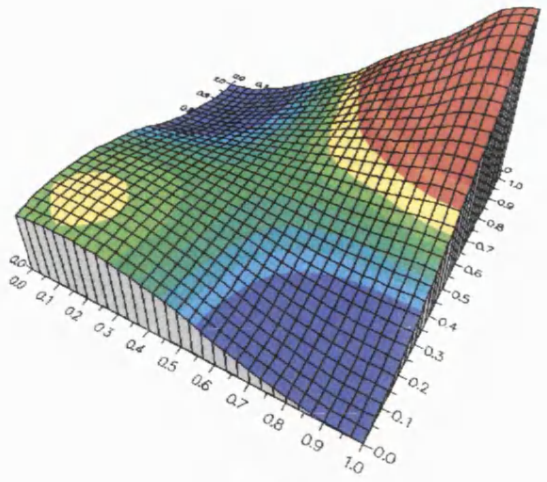
OH



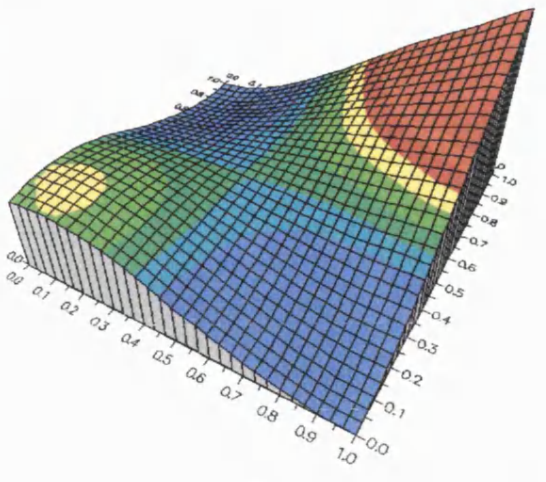
MeO



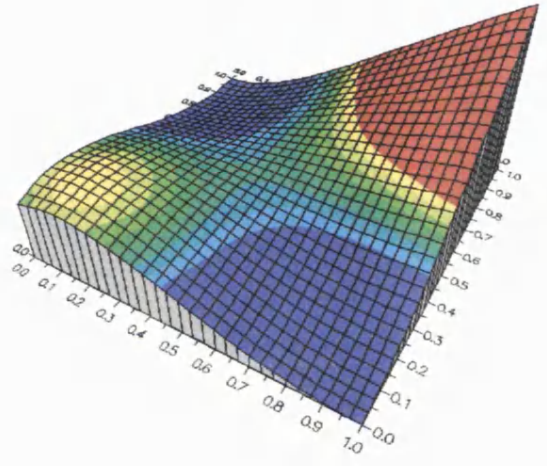
Me



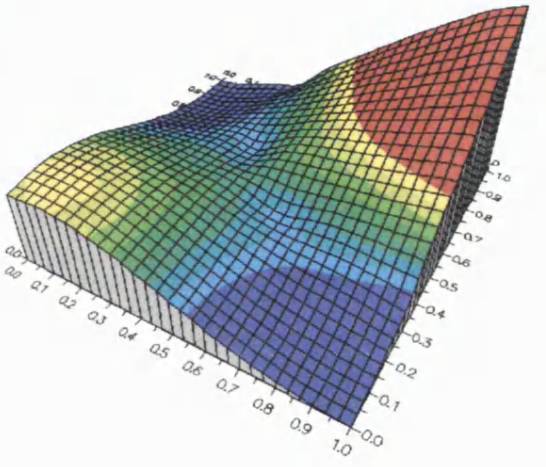
$CHCH_2$



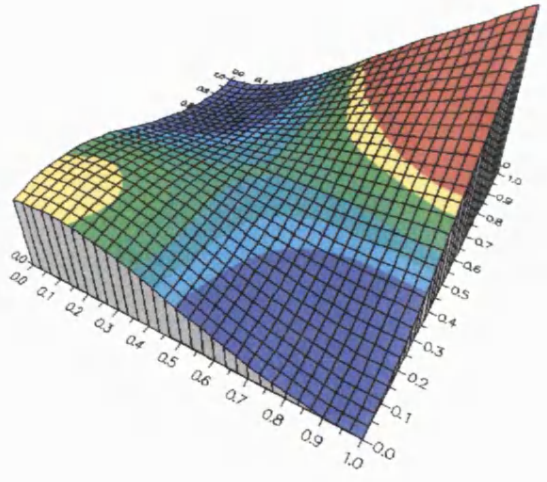
H



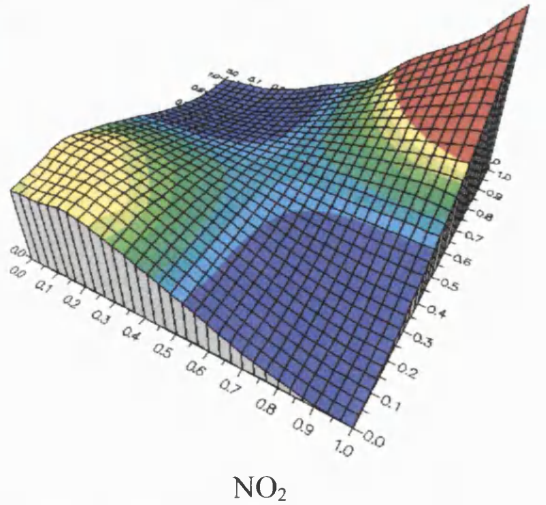
Br



CHO



CN



NO_2

Figure 5.2: AM1/COSMO generated energy surfaces depicting a gradual mechanistic changeover from S_N1 (X = NMe_2, NMeH, NH_2, OH, MeO) to S_N2 (X = $CHCH_2$, H, Br, CHO, CN, NO_2).

From the activation energies, the changeover of mechanism occurs between X=MeO, and X=Me substituents. Here, the difference in activation energies between the S_N2 and S_N1 mechanism is only 3.5 kJ mol^{-1} for X=Me. All of the surfaces show two distinct reaction pathways and none show a single, merging pathway. The X=NMe_2, NMeH, NH_2, OH and MeO *p*-substituted benzyl energy surfaces show that these substituents are sufficiently powerful electron donors for the S_N1 path to be preferred. On the other hand, X=$CHCH_2$, H, Br, CHO, CN and NO_2 substituted benzyl chlorides are not sufficiently stabilising to allow the S_N1 route to be viable and so the S_N2 route is favoured, Figure 5.2. The methyl substituted, X=Me, shows a borderline/intermediate case where both S_N1 and S_N2 mechanisms run concurrently differing only slightly in barrier heights. The surfaces before, at and past this transitional point show a 'see-saw' effect between the S_N1 to S_N2 reaction mechanisms with the methyl case showing an almost balanced situation between two coexisting, concurrent mechanisms.

5.4.4 Structure-activity relationship

The Hammett equation has often been used as a means of correlating structure and reactivity of benzylic substrates, and as a mechanistic tool to probe the properties of the transition state. Rate constants k_x for the reaction of a compound containing an aryl group with a substituent X in either the *meta* or *para* positions are correlated with equilibrium constants k'_x for ionisation of similarly substituted benzoic acids, according to Equation 5.5.

$$\log (k_X/k_H) = \log (k'_X/k'_H) \times \rho \qquad \text{Equation 5.5}$$

The logarithm of the ratio of equilibrium constants on the right hand side of this equation is simply the pK_a of the X-substituted benzoic acid (in water at 25°C) relative to the benzoic acid itself, which defines the substituent σ, characteristic of each substituent group X in a particular series (either *meta* or *para*), Equation 5.6.[69]

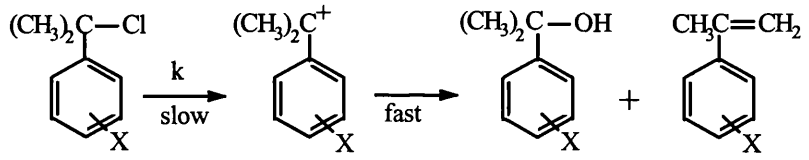
$$\sigma = (pK_a)_H - (pK_a)_X \qquad \text{Equation 5.6}$$

Electron-withdrawing substituents have positive values for σ, since they enhance the acidity of substituted benzoic acid, whereas electron-donating substituents have negative values of σ, since they diminish the acidity of substituted benzoic acid. Equation 5.5 is usually rewritten more simply as Equation 5.7.

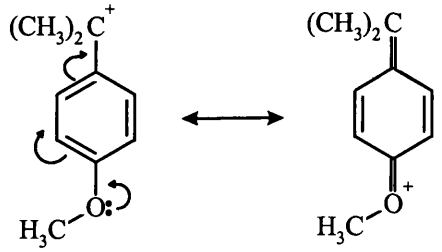
$$\log(k_X/k_H) = \sigma\rho \qquad \text{Equation 5.7}$$

The slope of a linear correlation between $\log(k_X/k_H)$ and σ is the reaction constant, which is a measure of the sensitivity of the reaction under consideration to the electronic influence of a substituent X, as compared with the influence of that substituent on benzoic acid ionisation.

The rate determining step for the S_N1 solvolysis of a series of substituted phenyldimethylcarbinyl chlorides is the heterolysis of the C–Cl bond leading to the formation of a carbocationic intermediate. Since the transition-state is carbocationic in nature, the influence of substituents upon the rate constant k is in the opposite direction to that for benzoic ionisation, in which a negative charge is being generated; hence ρ has a negative value for the reaction. However, it is observed that strongly electron-donating substituents in the *para* position give rate constants for the hydrolysis of phenyldimethylcarbinyl chlorides in 90% aqueous acetone at 25°C, Equation 5.8, that are larger than would be expected by their values according to the linear correlation of $\log(k_X/k_H)$ vs. σ for the other substituents. This is because the carbocationic intermediate, and the transition-state for its formation, can be stabilised more effectively by these substituents than is expected, owing to through-conjugation between the *p*-substituent and the benzylic position, Equation 5.9.



Equation 5.8



Equation 5.9

This resonance effect is not possible in the corresponding *p*-substituted benzoic acid. The points for the strongly electron-donating substituents in the *para* position may be restored to the linear Hammett correlation by use of a modified substituent constant σ^+, Figure 5.3. The demonstration of a better linear correlation for $\log(k_X/k_H)$ vs. σ^+ than for $\log(k_X/k_H)$ vs. σ provides evidence for the mechanistic involvement of a transition-state bearing positive charge at the benzylic position.

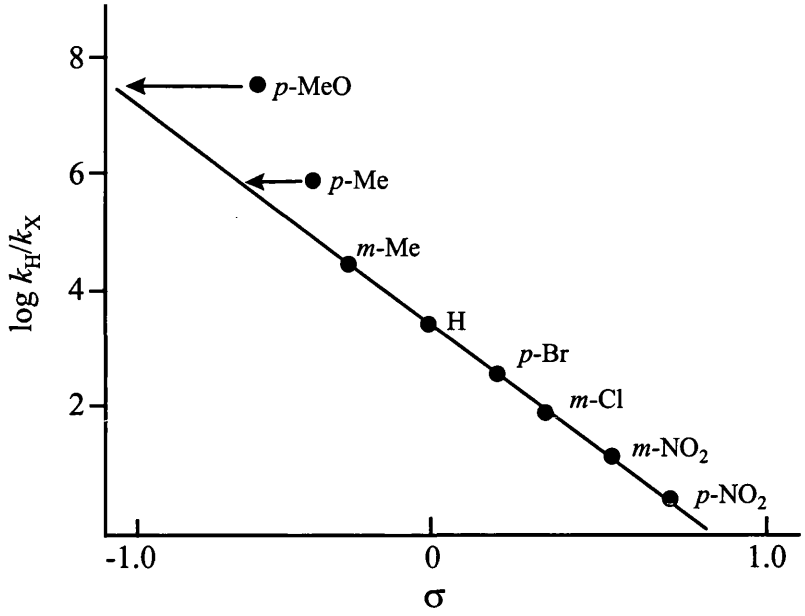


Figure 5.3: A Hammett plot vs. σ for the solvolysis of substituted phenyldimethyl-carbinyl chlorides in 90% aqueous acetone at 25°C.

When substituted benzyl substrates react with negatively charged nucleophiles, U-shaped Hammett plots are obtained. The minimum in the Hammett plot occurs in the vicinity of the unsubstituted substrate regardless of nucleophile, leaving group or solvent. The explanation for U-shaped or curved Hammett plots is that either there is a change in mechanism, from S_N2 to S_N1, or that there is a change in structure of the transition-state for a single mechanism with a differing balance of bond formation, such that electron-donating substituents cause a shift to a transition-state with more positive charge development on the central carbon. Another possible explanation is that different substituents act to stabilise the transition-state with a differing balance of polar and resonance effects (a two interaction mechanism).[70]

The linearity of a Hammett plot over a particular range of substituents is usually taken as evidence for a single rate limiting step for reactions involving a transition-state of essentially constant structure. Observation of Hammett relationships involving two intersecting straight lines or that possess significant curvature has been taken as evidence for a changeover in mechanism or a change in transition-state structure for a single mechanism. For the solvolysis of substituted benzyl tosylates in aqueous acetone Hammond *et al.* found that the data was best represented by a smooth curve for substituents ranging from *p*-methoxy to *p*-nitro.[71]

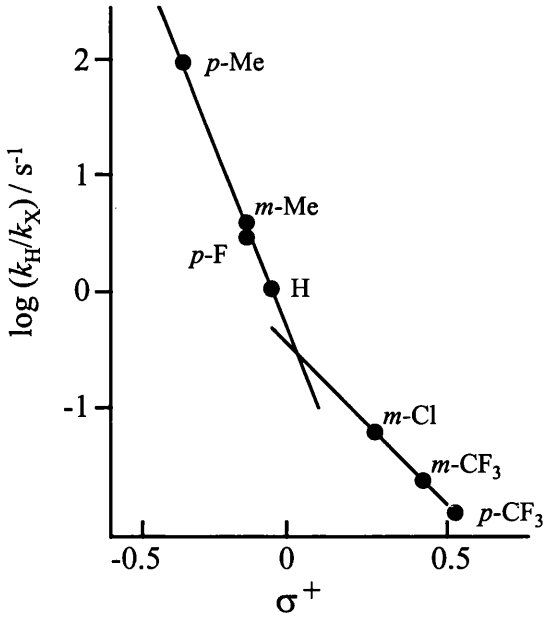


Figure 5.4: A Hammett correlation vs. σ^+ for the acetolysis of substituted benzyl *p*-toluene sulphonates showing two linear correlations.[30]

The acetolysis of substituted benzyl *p*-toluene sulphonates produced a curved Hammett correlation vs. σ^+, Figure 5.4. The reaction was said to occur by a single S_N2 mechanism that is 'pure' S_N2 when the substituent effect is electron-withdrawing. As the substituent is made more electron-donating, the S_N2 transition-state is stabilised and starts to resemble a carbocation. The S_N2 carbocation resembles a S_N1 transition-state however they still classed the mechanism as S_N2.[30]

The curvature of Hammett plots has also been attributed to changes in the amount of bond formation and bond rupture at the transition state. If the bond formation is greater than the bond rupture at the transition state, the α-carbon will be less positively charged at the transition-state as compared to the reactants which means that the reaction would be accelerated by electron-withdrawing substituents and subsequently ρ will be positive. On the other hand, if bond rupture is greater than the bond formation the α-carbon will be more positively charged at the transition-state as compared to the reactants which means that the reaction would be accelerated by electron-donating substituents and subsequently ρ will be negative. An important factor in determining transition-state structure is the relative bond strengths of the α-carbon-nucleophile and α-carbon-leaving group and the way in which the reacting bond is altered by a change in substituent. It was proposed by Westaway[66] that a change of substituent on the α-

carbon would effect the weaker bond but have considerably less or even no effect on the stronger bond. The introduction of a more electron-withdrawing substituent to the aryl ring shortens the α-carbon-nucleophile bond and reduces the conjugation between the α-carbon and the benzene ring but does not alter the α-carbon-leaving group distance, Figure 5.5.

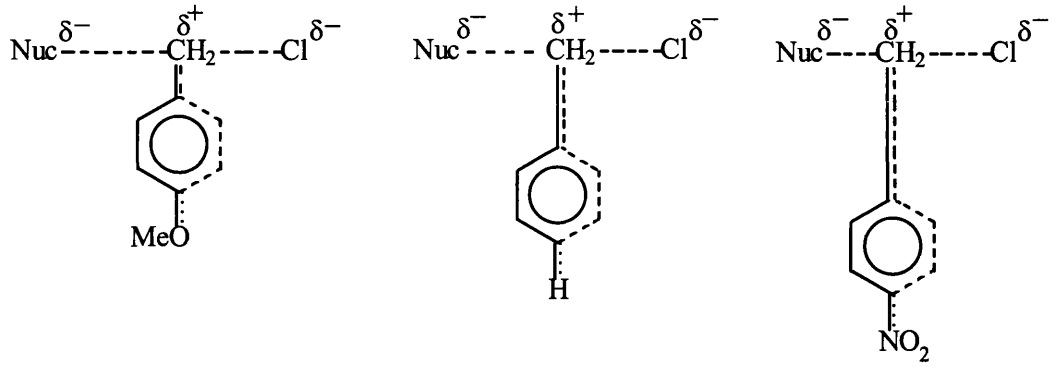


Figure 5.5: The effect of changing the *p*-substituent on the transition-state structure for S_N2 reactions.

The experimental Hammett correlation vs. σ for the reaction of carbanion nucleophiles with a family of N-benzyl-N,N-dimethyl anilinium cations and simply alkyl halides possess distinctively U-shaped plots. The U-shaped plot was said to originate due to the fact that the structure of the S_N2 transition-state were varying as the substituents were made to vary.[72] The bimolecular kinetic study of the benzylation of anilines in ethanol gave linear Hammett correlations vs. σ on varying the substituents on the aniline but showed considerable curvature on varying the benzyl chloride.[73] This reaction was said to occur by the S_N2 mechanism with a transition-state that varies in tightness as the benzyl substituent is varied.

Different substituents interact with the transition-state with differing balance of polar and resonance effects. Electron-donating substituents mainly stabilise a cationic transition-state through a resonance effect whereas electron-withdrawing substituents destabilise the transition-state through a polar effect. The Yukawa-Tsuno[74] equation allows for the varying relative amounts of electron donation by resonance to be correlated, Equation 5.10, where $\sigma^+ - \sigma$ is an empirical measure of the ability of a particular substituent to donate electrons by resonance and r^+ varies in some manner with the varying contributions of resonance effects in different reactions.

$$\log (k_H/k_X) = \rho(\sigma + r^+ (\sigma^+ - \sigma)) \qquad \text{Equation 5.10}$$

If r^+ is 0, then the equation simplifies to Equation 5.7 and when r^+ is equal to unity, we have the obedience of σ^+. Young and Jencks[70] have shown that by using a modified Yukawa-Tsuno equation, it is possible to separate out polar and resonance effects, Equation 5.11.

$$\log (k_H/k_X) = \rho\sigma + \rho' (\sigma^+ - \sigma)) \qquad \text{Equation 5.11}$$

The equation is based on separate ρ and ρ' parameters for the polar and resonance effects, respectively. The advantage of the modified Yukawa-Tsuno equation is that ρ' is a direct measure of the contribution of resonance effects to a particular reaction. Young and Jencks investigated the reaction between bisulphite and substituted acetophenones at 25°C and plotted the bimolecular rates against σ, Figure 5.6, which gave a U-shaped curve with a minimum at the unsubstituted acetophenone. The rate constants for *p*-NO_2, *m*-Br and the unsubstituted compound, which are assumed to possess no ability to interact by resonance, form a linear series against σ where the slope, ρ, is a measure of the polar effect of the substituent. The resonance contribution of substituents can then be evaluated from the deviation of the points from the linear correlation of σ. A correlation of the deviations against $(\sigma^+ - \sigma)$ gives a value of ρ for the resonance contribution.

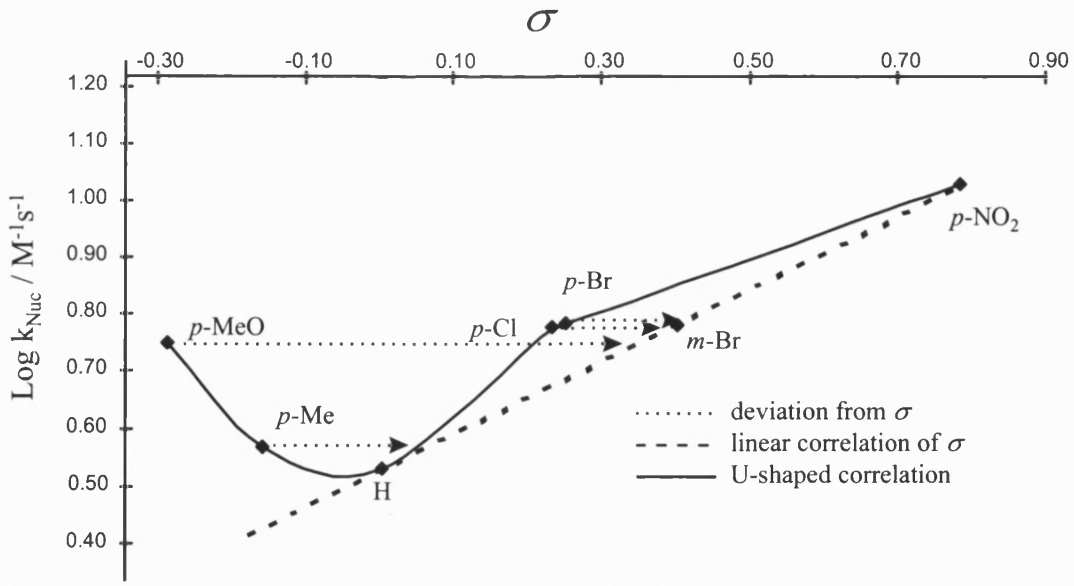


Figure 5.6: A Hammett plot vs. σ for the bimolecular reaction between bisulphite and substituted acetophenones at 25°C.

Therefore, in reactions on benzyl substrates with negatively charged nucleophiles, the upward curvature of the Hammett plot for electron-withdrawing substituents occurs because the transition-state are stabilised by electron withdrawal through the polar effect and the upward curvature of the Hammett plot for electron-donating substituents occurs because the transition-state are stabilised by electron donation through the resonance effect. The change between strong electron donation and strong electron withdrawal will occur at the unsubstituted benzyl chloride, as this where a minimum in rate will occur.

Hammett correlations are examples of linear free-energy relationships, since the logarithm of a rate constant (or an equilibrium constant) is proportional to a free energy of activation (or of reaction). Equation 5.7 may therefore be recast as

$$(\Delta G_H^{\ddagger} - \Delta G_X^{\ddagger}) = 2.303RT\ \sigma\rho \qquad \text{Equation 5.12}$$

which may in turn be expanded to:

$$(\Delta H_H^{\ddagger} - \Delta H_X^{\ddagger}) - T(\Delta S_H^{\ddagger} - \Delta S_X^{\ddagger}) = 2.303RT\ \sigma\rho \qquad \text{Equation 5.13}$$

For substituent X in the *para* position, the entropy of activation is approximately constant so that $(\Delta S_H^{\ddagger} - \Delta S_X^{\ddagger})$ approximates to 0. The Hammett equation can then be rewritten as:

$$-\Delta H_X^{\ddagger} = 2.303RT\ \sigma\rho - \Delta H_H^{\ddagger} \qquad \text{Equation 5.14}$$

Figure 5.7. shows a plot of the AM1/COSMO calculated enthalpies of activation (or rather the negatives of theses energy differences) against σ[75] (a plot against σ^{+} gave the same correlation as plotting against σ). The calculated results show that if a single mechanism is in operation, either the S_N2 or S_N1 mechanism, then the plot would be a single, linear line. In this case, two mechanisms are operating giving two separate linear portions that crossover at the methyl substituent. The larger magnitude of the Hammett plot for S_N1 than for S_N2 correlation is consistent with the larger change in charge for the S_N1 mechanism, Table 5.2. The α-carbon-benzene ring bond orders for the S_N2 mechanism are all small in magnitude across the series, Table 5.3. The *para* substituent effect does not have a major effect on the C_α-C_{AR} bond for the S_N2 mechanism. However, as expected, the C_α-C_{AR} bond for the S_N1 mechanism possess a great deal of double bond character at the transition-state. There is no evidence for a dramatic change in resonance effect of the substituents; the bond orders for the C_α-C_{AR} do not change

smoothly and predictability, but do not cause non-linearity in the S_N1 and S_N2 correlations.

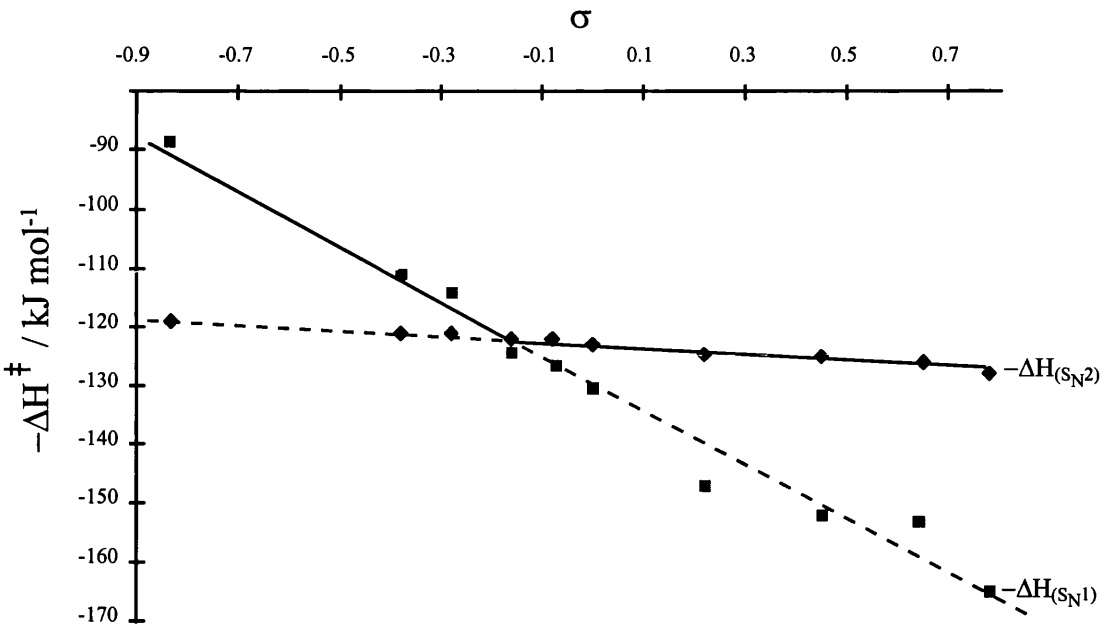


Figure 5.7: A Hammett plot vs. σ for the corresponding AM1/COSMO calculated energies, kJ mol^{-1}, S_N2 and S_N1 mechanisms.

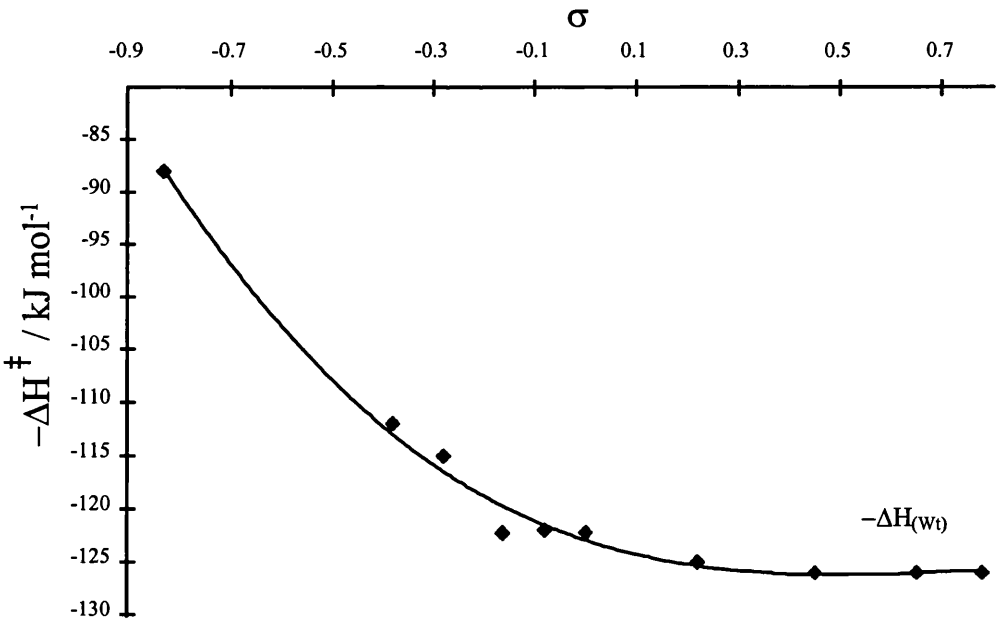


Figure 5.8: A Hammett plot vs. σ for the corresponding weighted barrier heights, Wt.

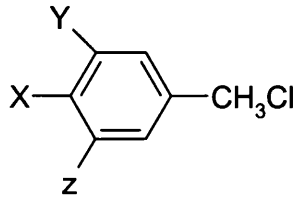
Chapter Six

Theoretical Modelling of S_N2/S_N1 Reactivity of Benzyl Derivatives: The Reaction of Substituted Benzyl Chlorides with Azide Anion

Chapter 6 - Theoretical Modelling of S_N2/S_N1 Reactivity of Benzyl Derivatives: The Reaction of Substituted Benzyl Chlorides with Azide Anion

6.1 Introduction

The logical extension to the investigation of the symmetric reaction of substituted benzyl chlorides is the investigation of an asymmetric reaction. An extensive study of the S_N2 and S_N1 reaction mechanisms of substituted benzyl chlorides is undertaken. The principal goals of this study is to characterise changes from stepwise to concerted reaction mechanisms for aliphatic nucleophilic substitution that occur as the carbocation intermediate of the stepwise reaction is destabilised by electron-withdrawing ring substituents. In conjunction with experimental data, it may be possible to determine the origin for curvature in the Hammett plots of rate data for these reactions and to characterise the changes in the transition-state structures that occur as the ring substituent is made electron-withdrawing relative to *p*-MeO.



X = MeO, Me, H, Br or NO_2
Y = MeO, H, Br, CN or NO_2
Z = H or NO_2

Equation 6.1

An experimental investigation was carried out by J. P. Richard on the solvolysis in 20% aqueous MeCN and nucleophilic substitution reactions at twenty-four ring substituted benzyl derivatives, Equation 6.1.[76] One of the aims of his investigation was to determine the cause of curvature that is present in certain Hammett plots of rate data for bimolecular nucleophilic substitution reactions, Figure 6.1(a), and for the solvolysis reactions at benzyl derivatives, Figure 6.1(b). The L-shaped Hammett plot curvature with either σ or σ^+ shows that the transition-state for the these reactions are greatly stabilised by electron-donating substituents relative to H and only weakly destabilised by electron-withdrawing substituents. None of the data shows severe U-shaped curvature due to the electron-donating substituents slowing down the reaction instead of speeding them up. The Hammett plots for the bimolecular nucleophilic substitution resembles that for solvolysis reaction. This implies that these reactions undergo similar changes in transition-state structure in going from electron-donating to withdrawing ring substituents.

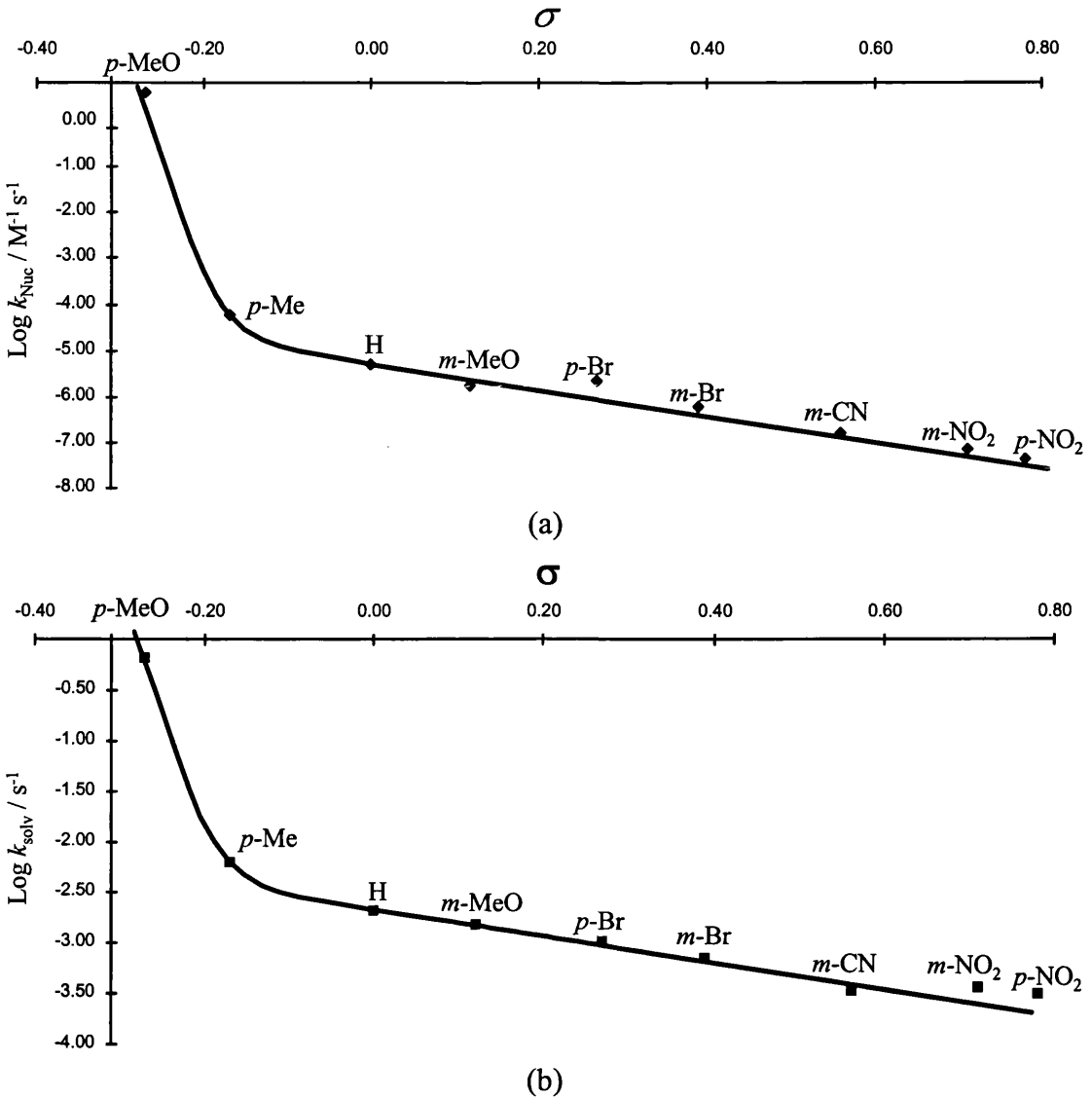


Figure 6.1: Hammett plots for the reaction of mono-substituted benzyl chlorides against σ for (a) the bimolecular nucleophilic substitution and (b) the solvolysis reaction in 20% aqueous MeCN at 25°C.

It is postulated that the substrates on the left hand side of the Hammett plot undergo reaction by a stepwise mechanism in which the rate limiting step is the trapping of a carbocation-leaving group ion-pair by the nucleophile. The stepwise reaction will proceed through a carbocation-like transition-state that will be stabilised by electron-donating ring substituents. Another possibility is that there is a change in mechanism from stepwise to concerted as the electron-withdrawing substituents destabilise the transition-state and the S_N1 intermediate well disappears forcing a bimolecular reaction, i.e. a change in transition-state structure for a single mechanism only. The curvature may also be due to the S_N1 and S_N2 mechanisms running concurrently with one

mechanism being dominant over the other. As the stability of the carbocation is increased, there is an increase in the amount attributed to the S_N1 pathway.

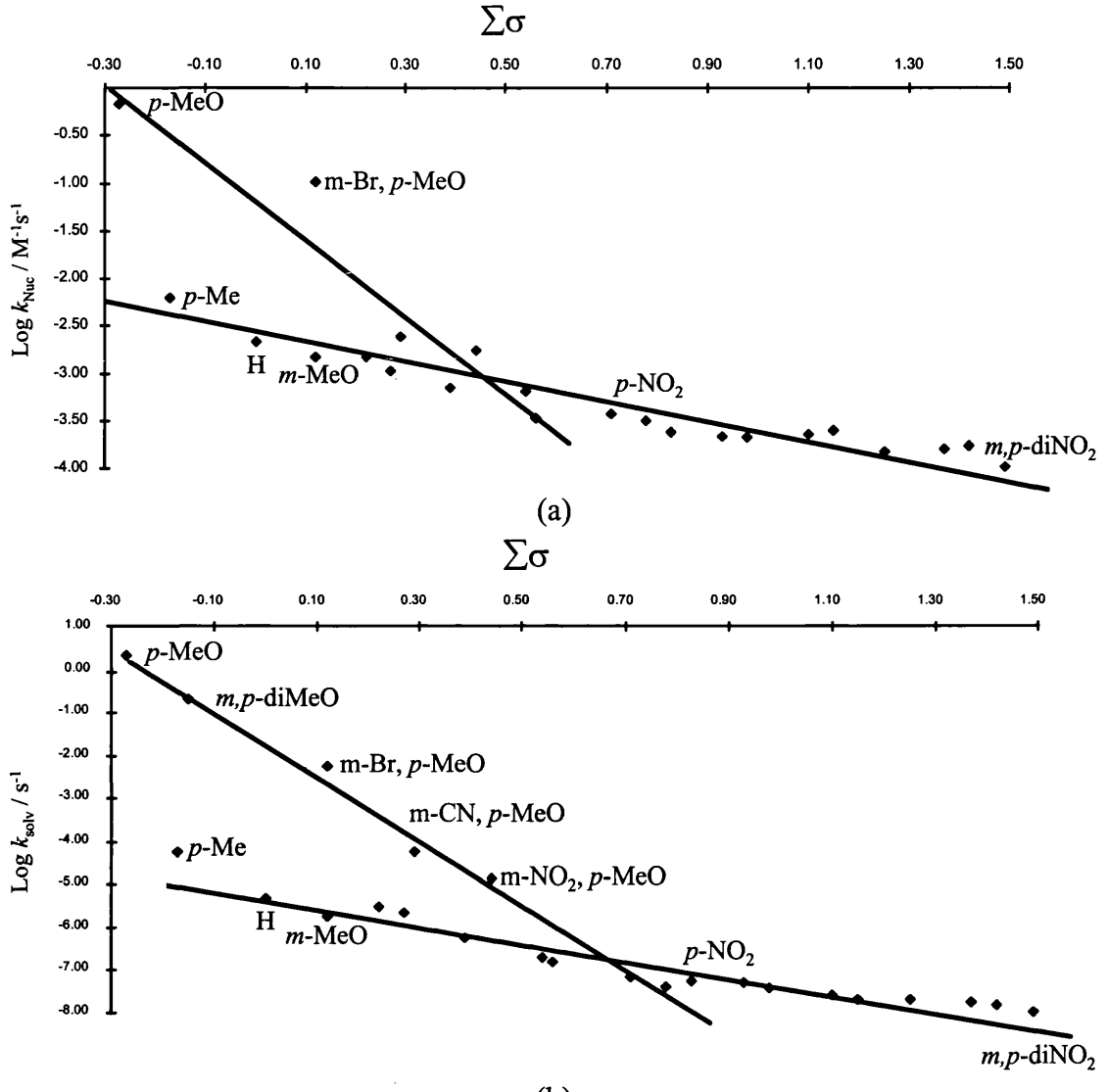
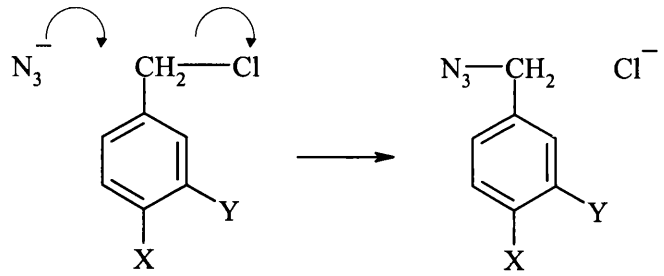


Figure 6.2: A Hammett plot for the reaction of the twenty-four different substituted benzyl chlorides against $\sum\sigma$ for (a) the bimolecular nucleophilic substitution and (b) the solvolysis reaction in 20% aqueous MeCN at 25°C.

A Hammett plot for the reaction of the twenty-four different substituted benzyl chlorides against σ for the bimolecular nucleophilic substitution, Figure 6.2(a), and for the solvolysis reaction, Figure 6.2(b), resemble each other. For the bimolecular series, only two points lie off the linear correlation; the *p*-methoxy benzyl chloride and *m*-bromo-*p*-methoxy benzyl chloride. For the solvolysis series of reactions, two linear correlations can be plotted through the data, one being steeper in magnitude than the

other. The steeper correlation is composed of substituents with the methoxy group in the *para* position.

6.2 Methods



X = MeO, Me, H, F, CF_3O or NO_2
Y = MeO, Me, H, Br, F, CN or NO_2

Figure 6.3: The *m*- and *p*-disubstituted benzyl chloride reaction with azide anion (N_3^-) where X and Y are the variable substituents.

Computational modelling provides a tool for the investigation of chemical reactivity. In order to complement the experimental work carried out by J. P. Richard, the S_N2 and S_N1 activation energies for twenty *m,p*-disubstituted benzyl chlorides is undertaken. The AM1 semi-empirical Hamiltonian utilising COSMO,[11] a continuum model, in MOPAC93[12] is used to model the reaction in water. The lowest energy orientation for the transition-states are obtained with the C–Cl and C–N bond vector perpendicular to the benzene ring. Such an orientation in the transition-state allows for the overlap of the p-orbitals on the chlorine and nitrogen with the π-system in the benzene ring which lowers the systems energy. Each of the transition-state structures is characterised as having only one imaginary frequency. The CAMVIB and CAMISO programs are used to compute the kinetic isotope effects at 298.15K.[28]

6.3 Results

The activation energy for the nucleophilic substitution reactions for the S_N2 and S_N1 mechanisms are calculated for the series of the different substituted benzyl chlorides, Table 6.1. The enthalpy of activation for the S_N2 mechanism is defined as

$$\Delta H^{\ddagger}_{(S_N2)} = \Delta H_{f(S_N2\ TS)} - \{\Delta H_{f(XYBzCl)} + \Delta H_{f(N_3^-)}\} \text{ / kJ mol}^{-1}. \qquad \text{Equation 6.1}$$

The enthalpy of activation for the S_N1 mechanism is defined as

$$\Delta H^{\ddagger}_{(S_N1)} = \Delta H_{f(S_N1\ TS)} - \Delta H_{f(XYBzCl)} \text{ / kJ mol}^{-1}. \qquad \text{Equation 6.2}$$

The transition-state bond lengths (Å) for the carbon-chlorine (C–Cl) and carbon-nitrogen (C–N) distances for the various mechanisms are given together with the

Pauling bond orders which are calculated in the same manner as described in Chapter 4. The calculated transition-state-reactant charge difference for the nucleophile, leaving and benzyl group at the transition-state are given in Table 6.2. This is defined as the total charge on the benzyl group (or nucleophile or leaving group) at the transition-state minus the total charge on the benzyl group (or nucleophile or leaving group) of the reactants. Table 6.3 contains the α-deuterium and carbon kinetic isotope effect at 298.15K for both mechanisms. The α-deuterium kinetic isotope effects, k_H/k_D, are given per deuterium atom. Table 6.4 contains the α-carbon-benzene ring carbon bond lengths, C_α-C_{AR}, for the reactants and at the transition-state for the various mechanisms together with Pauling bond orders. The proportionality constant $c = 0.3$ was used to calculate the Pauling double bond orders.

Table 6.1: AM1/COSMO calculated activation energies, kJ mol^{-1}, bond lengths and Pauling bond orders for the S_N2 and S_N1 mechanisms.

Substituent		S_N2					S_N1		
X	Y	$\Delta H^{\ddagger}$	C–Cl / Å	C–N / Å	n_{C-Cl}	n_{C-N}	$\Delta H^{\ddagger}$	C–Cl / Å	n_{C-Cl}
MeO	MeO	160.9	2.40	2.10	0.35	0.34	124.7	2.96	0.14
MeO	H	157.4	2.42	2.11	0.35	0.34	114.4	2.95	0.17
MeO	Br	162.7	2.40	2.10	0.36	0.35	128.4	3.02	0.13
MeO	NO_2	165.0	2.36	2.09	0.38	0.35	132.9	3.01	0.13
Me	Me	157.6	2.40	2.10	0.35	0.34	122.6	3.00	0.15
Me	H	159.5	2.41	2.10	0.35	0.34	125.3	3.01	0.16
Me	Br	162.2	2.37	2.09	0.37	0.34	135.3	3.06	0.12
Me	NO_2	165.2	2.35	2.08	0.39	0.35	143.6	3.00	0.13
H	H	160.5	2.39	2.10	0.36	0.34	130.6	3.10	0.14
H	Br	163.7	2.36	2.08	0.38	0.35	137.5	3.04	0.13
H	CN	164.6	2.36	2.08	0.38	0.35	144.6	3.01	0.13
H	NO_2	166.0	2.34	2.08	0.39	0.35	151.9	3.00	0.13
F	H	162.4	2.39	2.09	0.36	0.34	136.6	3.01	0.13
F	Br	167.8	2.36	2.08	0.38	0.35	140.9	3.00	0.13
F	NO_2	167.4	2.34	2.08	0.39	0.35	155.0	3.07	0.12
CF_3O	Br	163.9	2.33	2.07	0.40	0.36	145.1	2.95	0.14
CF_3O	NO_2	166.7	2.34	2.08	0.39	0.35	162.1	3.15	0.10
NO_2	Br	168.0	2.29	2.06	0.42	0.36	172.6	3.06	0.12
NO_2	CN	170.6	2.29	2.06	0.42	0.36	174.4	3.25	0.08
NO_2	NO_2	173.5	2.28	2.05	0.43	0.37	177.9	3.15	0.10

Table 6.2: AM1/COSMO charge distribution at the transition-state for the nucleophile, benzyl and leaving group for the S_N2 and S_N1 mechanisms.

Substituent		S_N2			S_N1	
X	Y	$\Delta q_{N_3^-}$	Δq_{benzyl}	Δq_{Cl}	Δq_{Cl}	Δq_{benzyl}
MeO	MeO	-0.22	0.47	-0.69	-0.79	0.79
MeO	H	-0.22	0.46	-0.68	-0.78	0.78
MeO	Br	-0.22	0.47	-0.69	-0.80	0.80
MeO	NO_2	-0.22	0.47	-0.69	-0.81	0.81
Me	Me	-0.22	0.47	-0.69	-0.81	0.81
Me	H	-0.22	0.47	-0.69	-0.81	0.81
Me	Br	-0.23	0.45	-0.68	-0.81	0.81
Me	NO_2	-0.23	0.44	-0.67	-0.81	0.81
H	H	-0.22	0.46	-0.68	-0.80	0.80
H	Br	-0.23	0.44	-0.67	-0.81	0.81
H	CN	-0.23	0.44	-0.67	-0.80	0.80
H	NO_2	-0.24	0.42	-0.66	-0.81	0.81
F	H	-0.22	0.46	-0.68	-0.80	0.80
F	Br	-0.23	0.44	-0.67	-0.81	0.81
F	NO_2	-0.24	0.42	-0.66	-0.83	0.83
CF_3O	Br	-0.25	0.40	-0.65	-0.81	0.81
CF_3O	NO_2	-0.24	0.43	-0.66	-0.82	0.82
NO_2	Br	-0.25	0.40	-0.65	-0.83	0.83
NO_2	CN	-0.25	0.40	-0.65	-0.84	0.84
NO_2	NO_2	-0.25	0.39	-0.64	-0.84	0.84

Table 6.3: AM1/COSMO calculated α-deuterium, k_H/k_D, and primary carbon kinetic isotope effects for the S_N2 and S_N1 mechanisms.

Substituent		S_N2		S_N1	
X	Y	k_H/k_D	$k_{C^{12}}/k_{C^{14}}$	k_H/k_D	$k_{C^{12}}/k_{C^{14}}$
MeO	MeO	1.031	1.004	1.162	1.002
MeO	H	1.055	1.002	1.160	1.001
MeO	Br	1.031	1.004	1.159	1.003
MeO	NO_2	1.031	1.003	1.171	1.004
Me	Me	1.053	1.002	1.150	1.003
Me	H	1.053	1.002	1.167	1.002
Me	Br	1.028	1.002	1.171	1.004
Me	NO_2	1.027	1.002	1.167	1.002
H	H	1.045	1.002	1.189	1.004
H	Br	1.016	1.001	1.172	1.004
H	CN	1.003	0.989	1.181	1.005
H	NO_2	1.024	1.000	1.182	1.007
F	H	1.023	1.003	1.190	1.007
F	Br	1.008	1.001	1.175	1.005
F	NO_2	1.014	1.000	1.190	1.007
CF_3O	Br	1.004	1.000	1.182	1.004
CF_3O	NO_2	1.000	1.000	1.195	1.004
NO_2	Br	1.001	0.996	1.210	1.004
NO_2	CN	1.000	0.995	1.207	1.007
NO_2	NO_2	0.995	0.995	1.201	1.001

Table 6.4: AM1/COSMO calculated bond lengths and bond orders for the α-carbon-benzene ring carbon bond, C_α-C_{AR}, for the S_N2 and S_N1 mechanisms.

Substituent		Reactant	S_N2	S_N1	S_N2	S_N1
X	Y	C_α-C_{AR}/Å	C_α-C_{AR}/Å	C_α-C_{AR}/Å	$n_{C_\alpha\text{-}C_{AR}}$	$n_{C_\alpha\text{-}C_{AR}}$
MeO	MeO	1.481	1.444	1.378	1.13	1.41
MeO	H	1.482	1.443	1.378	1.14	1.41
MeO	Br	1.483	1.442	1.379	1.14	1.41
MeO	NO_2	1.481	1.441	1.379	1.14	1.40
Me	Me	1.480	1.445	1.378	1.12	1.40
Me	H	1.480	1.446	1.378	1.12	1.40
Me	Br	1.481	1.448	1.382	1.11	1.39
Me	NO_2	1.483	1.452	1.391	1.11	1.36
H	H	1.481	1.446	1.387	1.12	1.37
H	Br	1.482	1.451	1.388	1.11	1.37
H	CN	1.481	1.451	1.388	1.11	1.37
H	NO_2	1.482	1.454	1.392	1.10	1.35
F	H	1.481	1.447	1.380	1.12	1.40
F	Br	1.481	1.451	1.383	1.11	1.39
F	NO_2	1.481	1.454	1.384	1.10	1.38
CF_3O	Br	1.482	1.454	1.385	1.10	1.38
CF_3O	NO_2	1.482	1.454	1.387	1.10	1.37
NO_2	Br	1.483	1.461	1.401	1.08	1.31
NO_2	CN	1.485	1.462	1.400	1.08	1.32
NO_2	NO_2	1.484	1.464	1.403	1.07	1.31

6.4 Discussion

In order to investigate the influence of substituents at the *meta* and *para* positions of the benzyl ring on S_N2 and S_N1 mechanisms, their activation energies are calculated, Table 6.1. The results show that the S_N2 transition-states are stabilised by electron-donating substituents in either the *m*- or *p*- positions, and slightly destabilised by electron-withdrawing groups. However there is little overall variance in activation energy across the series. The *m,p*-dinitro substituted benzyl chloride possesses the highest activation energy, 173.5 kJ mol^{-1}, while the *p*-methoxy derivative possesses the lowest, 157.4 kJ mol^{-1}. As the substituents are made more electron-withdrawing, the transition-states become tighter as nucleophilic and leaving group stabilisation is required. The *p*-methoxy derivatives posses the loosest transition-state while the *p*-nitro derivatives posses the tightest. In going from electron-donating to withdrawing substituents, there is an increase in the transition-state bond order which is reflected in the decrease of Δq for the nucleophile, leaving and benzyl group. Neither *meta* or *para* substituents have a major effect on the C_α-C_{AR} bond for the S_N2 mechanism.

The kinetic isotope effects (KIE) for the substitution of a single deuterium for a proton at the α-carbon and the carbon isotope effects are calculated at 298.15K, Table 6.3. The S_N2 reaction mechanism proceeds through a transition-state in which the nucleophile and leaving group are both partially bonded to the benzylic carbocation. The benzyl groups with the more electron-donating substituents have looser transition-states and so have a greater α-deuterium KIE than benzyl groups with electron-withdrawing substituents. This relationship can be seen with the methoxy substituted benzyl families which have the loosest transition-states and the largest α-deuterium KIE whereas *m,p*-dinitro benzyl chloride has the tightest transition-state and the only inverse α-deuterium KIE. The carbon isotope effects are not inline with prediction as the calculated results show that the tightest transition-states give inverse values.

The S_N1 reaction mechanism proceeds through a carbocation transition-state that is greatly stabilised by electron-donating groups which interact with the developing positive charge. This is reflected in the lower activation energies for benzyl derivatives with electron-donating substituents. As the electron-donating ability of substituent is replaced by electron-withdrawing substituents, the carbocation stability is reduced. Due to this decrease in carbocation stability, the activation energy increases. This manifests itself in a difference of approximately 40 kJ mol^{-1} between the *p*-methoxy and *p*-nitro benzyl families. The greater looseness of the S_N1 transition state as compared to the S_N2 is reflected by a greater Δq value for the benzyl and leaving group. The C_α-C_{AR} bond for the S_N1 mechanism possess a great deal of double bond character at the transition-state. As the nature of the electron-withdrawing substituents is increased, the bond between carbocation and leaving group increases due to the destabilisation of the carbocation. This in turn leads to a looser transition-state which is associated with a large α-deuterium KIE. This relationship can be seen with the *p*-methoxy family which have the tightest transition-states and the smallest α-deuterium KIE, whereas the *p*-nitro family possess the loosest transition-state and greatest α-deuterium KIE.

6.4.1 Structure-activity relationships

The AM1/COSMO calculated energies for the S_N2 and S_N1 mechanisms for the reaction with mono-substituted benzyl chlorides are plotted against the summation of Hammett σ values[75] (or rather the negatives of theses energy differences), Figure 6.3(a) and (b). The curved Hammett plots for the S_N2 mechanism resembles that for S_N1 reaction. This is because the reactions undergo similar changes in transition-state structure and a decrease in reactivity going from electron-donating to withdrawing ring substituents. The transition-states on the left hand side of benzyl chloride, σ = 0, are loose with a great deal of carbocation character for both the S_N1 and S_N2 mechanisms.

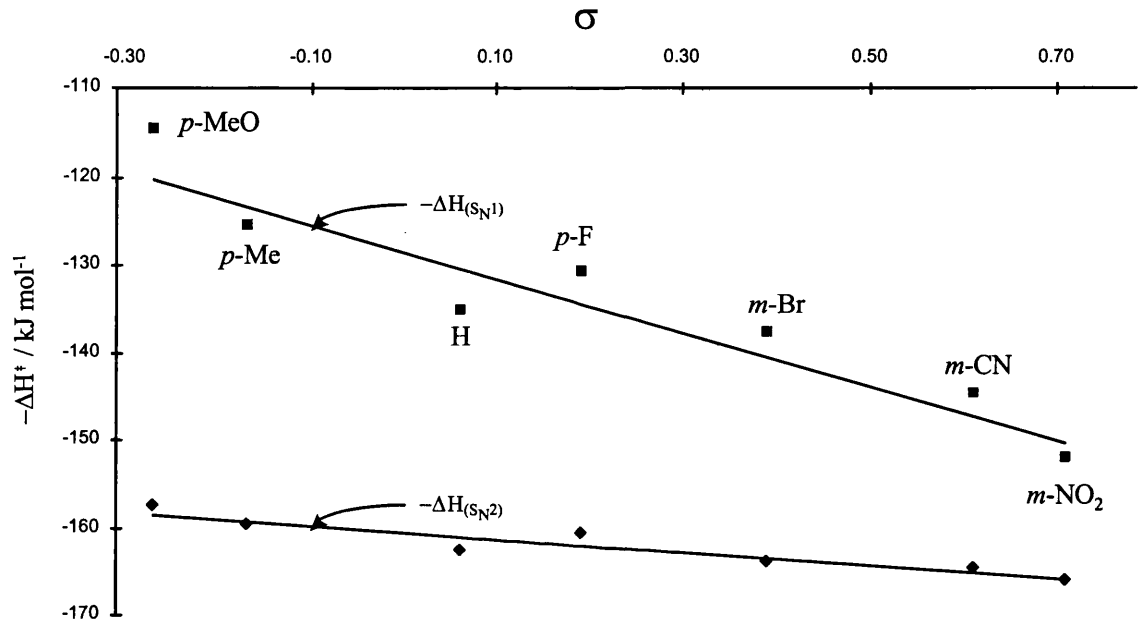


Figure 6.3: A Hammett plot for the reaction of mono-substituted benzyl chlorides against σ for the S_N2 and the S_N1 reaction mechanisms.

A Hammett plot for the reaction of the twenty different substituted benzyl chlorides against the σ for the S_N2 and S_N1 activation energies, Figure 6.4, do not resemble each other. A linear correlation can be plotted through the data for both the S_N2 and S_N1 series of reactions. For the calculated results, the crossover of the two linear correlations occurs when the substituents are strongly electron-withdrawing, which destabilise the carbocation of the S_N1 mechanism, and so enforces the S_N2 mechanism. The crossover point of the two linear correlations for the calculated results would shift to the left if the solvent modelled was less polar than water as a decrease in polarity of the solvent decreases the rate of the S_N1 mechanism. The crossover point of the two linear

correlations for the calculated results would shift to right if the reaction entropy would effect the rate of the S_N2 mechanism greater than it would effect the S_N1 mechanism.

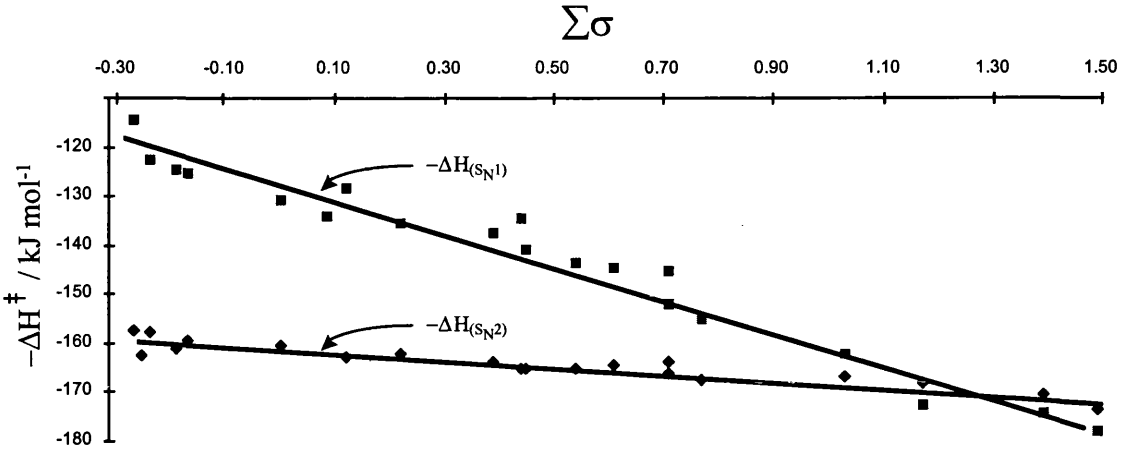


Figure 6.4: A Hammett plot for the reaction of twenty substituted benzyl chlorides against $\Sigma\sigma$ for the ring substituents for the S_N2 and the S_N1 reaction mechanisms.

Comparing the Hammett correlations for the experimental solvolysis data, Figure 6.2b with the calculated plot, Figure 6.4, an obvious similarity is noticed; both contain two linear correlations through the data. The shallower, lower linear correlation is due to the S_N2 mechanism while the steeper linear correlation is due to the S_N1 mechanism. From the experimental bimolecular data. The points that lie on the correlation, such as *m*-methoxy benzyl chloride, would react by a 'pure' S_N2 mechanism while the points that lie off the correlation, such as *p*-methoxy benzyl chloride, would react by a stepwise mechanism.

6.5 Conclusion

It has been demonstrated for the solvolysis reaction that linear correlations originate from a single mechanism operating. For solvolysis reactions, where the reaction is pseudo-first-order, two separate linear portions occur, one relating to the S_N2 mechanism being dominant, and the other steeper portion relating to the S_N1 mechanism. The L-shaped curvature for the solvolysis reaction is due to the S_N1 and S_N2 mechanisms running concurrently with one mechanism being dominant over the other. As the stability of the carbocation is increases, there is an increase in the amount attributed to the S_N1 pathway. The calculations performed do not reveal the cause of the departure from linearity for the bimolecular data, Figure 6.2(a), for points that lie off the linear correlation, such as *p*-methoxy benzyl chloride. It has been postulated that

substrates such as *p*-methoxy benzyl chloride undergo the reaction by a stepwise mechanism in which the rate limiting step is the trapping of a carbocation-leaving group ion-pair by the nucleophile. A quantum mechanical/molecular mechanical (QM/MM) model based on the default AM1 model in CHARMM[15] with transition-state refinement controlled by an external program, GRACE,[60] could be used to investigate the energetics of different rate limiting steps including the trapping of a carbocation-leaving group ion-pair.

Chapter Seven

The Transfer of the Methoxycarbonyl Group Between Isoquinoline and Substituted Pyridines in Aqueous Solution

Chapter 7- The Transfer of the Methoxycarbonyl Group between Isoquinoline and Substituted Pyridines in Aqueous Solution

7.1 Introduction

To probe the nature of activated-complex structure and properties is important because of the fundamental role of the activated-complex within the transition-state theory of chemical reactivity. Experimental mechanistic investigations yield results which are generally interrupted in terms of activated-complex structure; for example a non-linear Brønsted correlation indicates an intermediate borderline between concerted and stepwise mechanisms. It is usually difficult to predict the position of this breakpoint in a two-step mechanism. The exception is when the entering and leaving groups have similar structures. Assuming that a Brønsted dependence is under investigation for a stepwise mechanism involving the nucleophilic (Nu) displacement of the leaving group (Lg), Equation 7.1, then each step will have its own Brønsted equation ($\log k_{Nu} = \beta_n.pK_a + C_n$) and the measured rate constant will be governed by Equation 7.2 where k_0 is a constant, $\Delta pK = pK_{Nu} - pK_{Lg}$ and $\Delta\beta = \beta_2 - \beta_{-1}$. The β terms refer to the exponents for Brønsted-type equations for the individual rate constants in the scheme.

$$\text{A-Lg} + \text{Nu} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \text{Nu-A-Lg} \xrightarrow{k_2} \text{A-Nu} + \text{Lg} \qquad \text{Equation 7.1}$$

$$k_{obs} = k_0\, 10^{\beta_1.\Delta pK} / (1 + 10^{-\Delta\beta.\Delta pK}) \qquad \text{Equation 7.2}$$

Equation 7.2 predicts a free energy relationship with two straight lines intersecting at $\Delta pK = 0$, i.e. $k_{-1} = k_2$, when the change in rate limiting step occurs. A Brønsted plot must possess sufficient data points spread over a wide range of substitutional parameters to enable a reasonable statistical estimation of the slope.[77]

The observation of a Brønsted relationship involving two intersecting straight lines is taken as evidence for a change in the rate limiting step. The linearity of a free energy relationship is taken as evidence for a single limiting mechanism within the range of substituent parameters measured such as pK_a. Therefore a break in a plot would suggest a change in mechanism. A Brønsted plot can be seen for the attack of various substituted pyridines on 2,4,6-trinitrophenyl acetate,[31] Figure 7.1. The plot comprised

of two straight lines that intersected. The presence of this break point between two linear portions of a free energy relationship is regarded as evidence for the existence of the formation or breakdown of an intermediate in a stepwise mechanism.

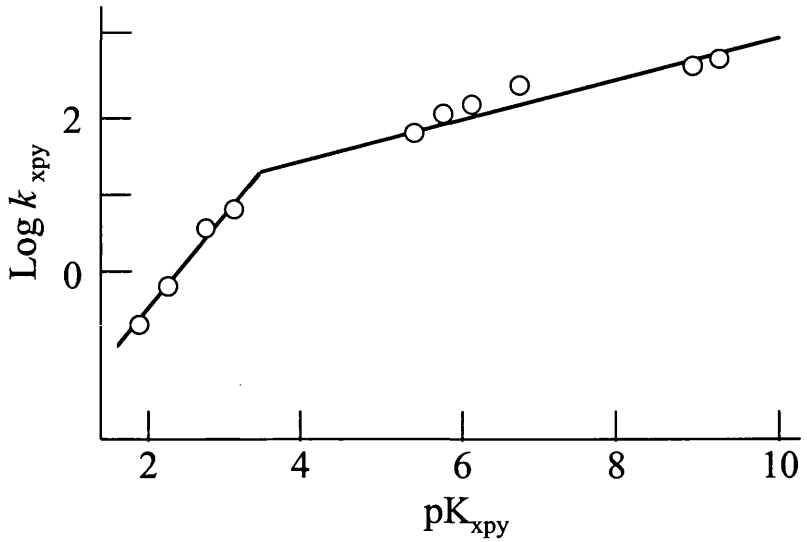


Figure 7.1: A Brønsted plot for the attack of substituted pyridines on 2,4,6-trinitrophenyl acetate indicating a stepwise mechanism.

An experimental investigation was carried out by Chrystiuk and Williams[78] on transfer of the methoxycarbonyl group between isoquinoline and various pyridine nucleophiles, Equation 7.3.

Equation 7.3

The study of acyl group transfer reaction in aqueous solution was undertaken using polar substituent effects to discriminate between a concerted and stepwise mechanism. The methoxycarbonyl group was chosen as it offered a reduction in rate compared to other functional groups enabling the kinetic to be followed more easily. Isoquinoline was employed as a leaving group because the spectral change enables the reaction to be followed at relatively low concentrations. The above considerations are not factors when following the reaction computationally. In their study, Chrystiuk and Williams came to the conclusion that the reaction was concerted and not stepwise in nature due to the fact that their Brønsted plot followed a good linear relationship over a wide pK_a range, Figure 7.2. A two step mechanism with two transition-states requires a 'break' in the Brønsted plot at the pK_a of isoquinoline. They stated that there was no statistical justification for this 'break', so a one step, concerted mechanism was proposed.

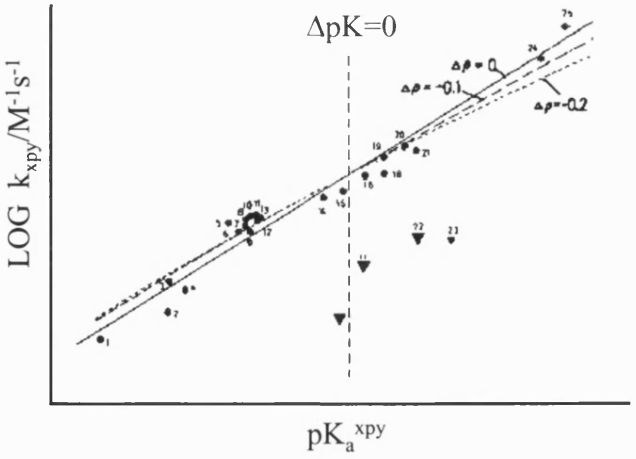


Figure 7.2: Chrystiuk and Williams' Brønsted plot for the transfer of the methoxycarbonyl group between isoquinoline and various pyridine nucleophiles.

7.2 Mechanistic detail

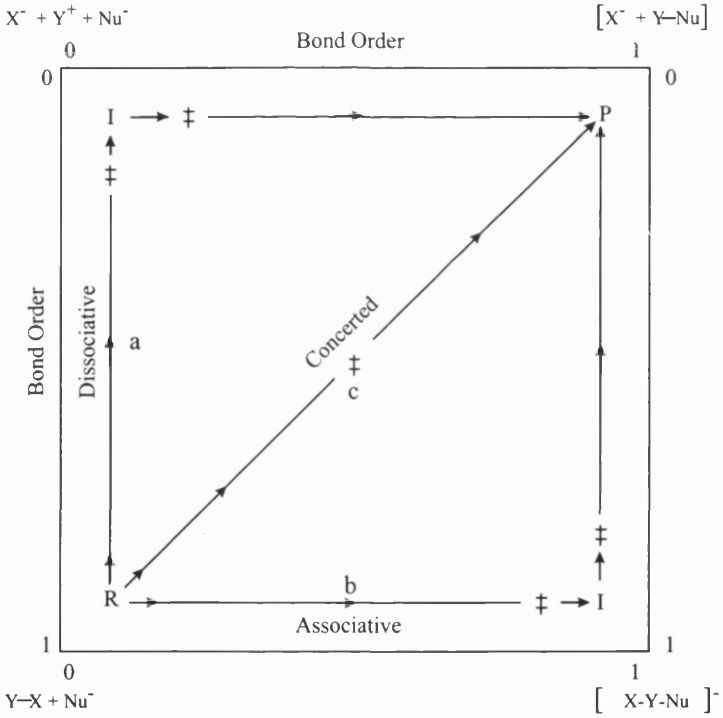


Figure 7.3: More O'Ferrall-Jencks energy surface diagram depicting various mechanistic possibilities for the transfer of a carbonyl group where R, P, ‡ and I represent the reactants, products, transition-states and intermediates respectively.

There are three possible mechanisms for the transfer of the carbonyl group between reactants and products. If Y is defined as the carbonyl group and X as the leaving group which will be substituted by the incoming nucleophile, Nu, then the various

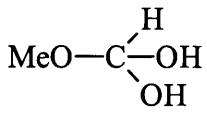
mechanisms; dissociative (a), associative (b) and concerted (c) can be summarised by the use of a More O'Ferrall-Jencks[43,44] diagram, Figure 7.3.

$$Y{-}X \xrightarrow{-X^-} Y^+ \xrightarrow{+Nu} Y{-}Nu \qquad \text{Equation 7.4}$$

The dissociative mechanism involves the ionisation of the reactants, YX, which is then attacked by the nucleophile, Nu, to form the products, Y-Nu, Equation 7.4. The acylium ion (RCO^+) is known to be stable in the gas phase and is very reactive in nucleophilic solvents. The stability of the acylium ion depends on 'internal stabilisation' and by the substrate possessing a good leaving group. The structure of the carbonyl group can be perturbed by the rest of molecule in which it resides. Both inductive and resonance effects can alter the electronic distribution in the carbonyl group therefore effecting the reactivity of the system. A double bond conjugated with the carbonyl group is a resonance stabilised system in which the carbonyl group is less reactive. The occurrence of an acylium ion intermediate in solution depends on the factors stabilising it and the stability of the leaving group.

$$Y{-}X \xrightarrow{+Nu^-} [X{-}Y{-}Nu]^- \xrightarrow{-X^-} Y{-}Nu \qquad \text{Equation 7.5}$$

The associative mechanism for the carbonyl group transfer reaction involves the formation of a tetrahedral intermediate in which the bond to the incoming group is made before the bond to the leaving group is broken, Equation 7.5. The conversion of the reaction centre from the trigonal to tetrahedral configuration will occur as the nucleophile forms a covalent bond to the carbon atom. This occurs because the carbon-oxygen double bond is weaker than the carbon-leaving group bond. The 'new' tetrahedral intermediate will possess a large amount of negative electron density centred on the oxygen therefore solvation of the reaction helps to stabilise the intermediate.



Equation 7.6

The problem with detecting these tetrahedral intermediates is their rapid decomposition to reactants or products. However certain tetrahedral intermediates have been isolated, Equation 7.6.[79]

$$Y{-}X \xrightarrow{+Nu^-} \left| X{-}Y{-}Nu \right|^{\ddagger -} \xrightarrow{-X^-} Y{-}Nu \qquad \text{Equation 7.7}$$

A concerted mechanism will have no intermediates thus will possess only a single transition-state, $[X\text{-}Y\text{-}Nu]^{\ddagger -}$, Equation 7.7.[80] It was thought that carbonyl group transfer

would not take place by concerted mechanisms as the associative mechanism, possessing a tetrahedral intermediate, was thought to be favoured.[81] This is due to the fact that a reaction involving the direct displacement of the leaving group as the nucleophile attacks, totally ignores the chemical properties of carbonyl group; that of addition and the ease of which it can form adducts. A possible transition-state for the concerted mechanism has a square planar geometry which would become tetrahedral, depending on the advancement of bond formation and fission to the entering and leaving group respectively.

A definition of an 'enforced' concerted mechanism can be thought of as involving a process passing through a structure corresponding to a species which is an intermediate but has a negligible barrier to decomposition, Figure 7.4.[82] This enforced concerted mechanism (B) could be thought as evolving from a changing condition in a stepwise mechanism (A), where the intermediate becomes so unstable its lifetime of decomposition is as short as that for a bond vibration (c.a. 10^{-3}s). The enforced concerted mechanism is a possible borderline between stepwise and a regular concerted mechanism (C).

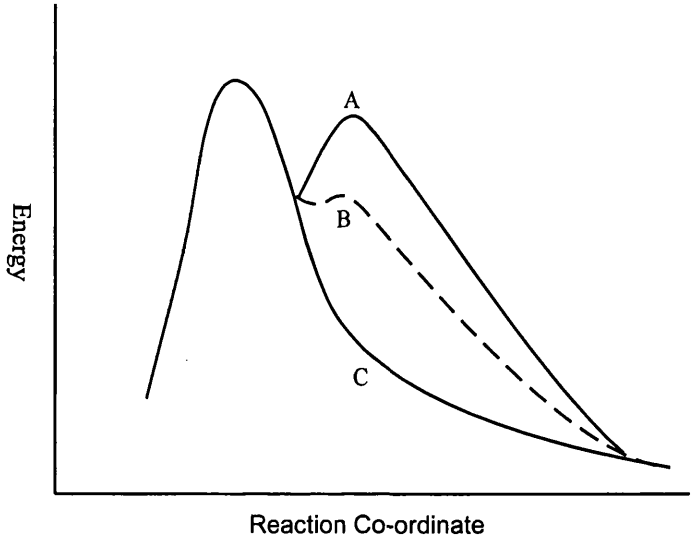


Figure 7.4: A two dimensional energy diagram illustrating a stepwise (A), enforced concerted (B) and a regular concerted mechanism (C).

7.3 Previous work

A theoretical gas-phase simulation of a rate-equilibrium correlation for the methoxycarbonyl group transfer between isoquinoline and a range of substituted pyridines, Equation 7.3, was carried out by Hammond and Williams.[83,84] An experimental Brønsted correlation had been published for a reaction series in aqueous solution, Figure 7.2. The presence of a linear correlation between log *k* (the logarithm of the observed second-order rate constant) and pK_a (of the conjugate acid of the nucleophile) over a wide range of basicity and the absence of a break point at the pK_a of isoquinoline were taken as evidence that the reaction proceeds in a concerted fashion rather than the usual stepwise addition-elimination mechanism for acyl transfers, Figure 7.5.

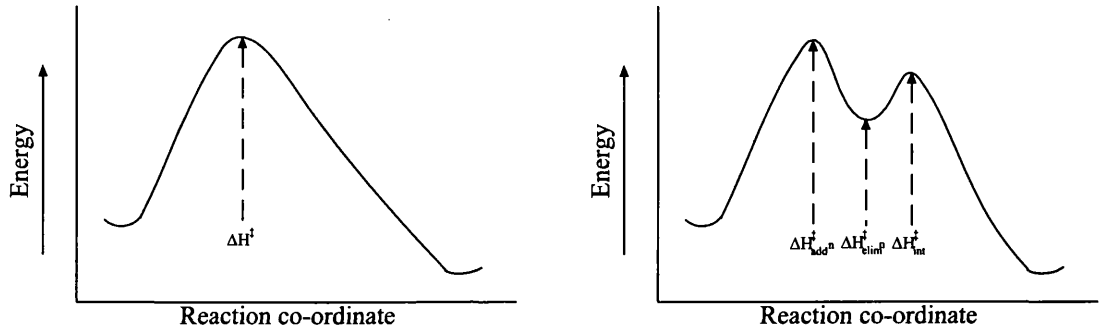
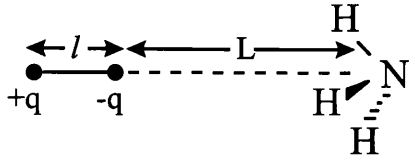


Figure 7.5: Energy profile diagrams for (a) S_N2 acyl transfer and (b) addition-elimination mechanism for acyl transfer.

In the study by Hammond and Williams, the pyridine nucleophiles and isoquinoline leaving group were mimicked by ammonia molecules whose proton affinity were modulated to reproduce the experimental proton affinity of each base relative to that of ammonia. Mimicking was accomplished by placing a pair of oppositely charged dipoles on the 3-fold axis of ammonia, on the same side as the hydrogens, with the negative charge at a constant distance L = 5.0Å from the nitrogen atom, as shown in Equation 7.8.



Equation 7.8

By adjusting the charge (|q| = 1 or 2) and the dipole separation (*l*) the basicity of the ammonia could be tuned. The proton affinities were calculated as the negative heats of

reaction, Equation 7.9, where q^-q^+ indicates a dipole moment and location which were the same in the reactant and product species.

$$NH_3.\ q^-q^+ + H^+ \rightarrow NH_4^+.\ q^-q^+ \qquad \text{Equation 7.9}$$

It was assumed that the contributions from the entropy were essentially constant and would not sensibly affect the relative values of $\Delta H^{\ddagger}$ and ΔH. This simplification allows the rate-equilibrium correlation to be constructed as a plot of $\Delta H^{\ddagger}$ vs. ΔH only. In practice the plot of the activation enthalpies were plotted against relative proton affinities, Figure 7.6. The plot comprised of two straight lines intersecting at the proton affinity of isoquinoline. The presence of a break point between two linear portions of a free energy relationship is regarded as evidence for the existence of the formation or breakdown of an intermediate in a stepwise mechanism. The advantage of the computaional study is that it allowed the characterisation of all the steps in the reaction profile, i.e. the barrier heights for the addition and elimination steps and the energy of the intermediate whereas experimentally only the rate determining step is seen. This allowed the generation of two complete intersecting lines in the Brønsted plot.

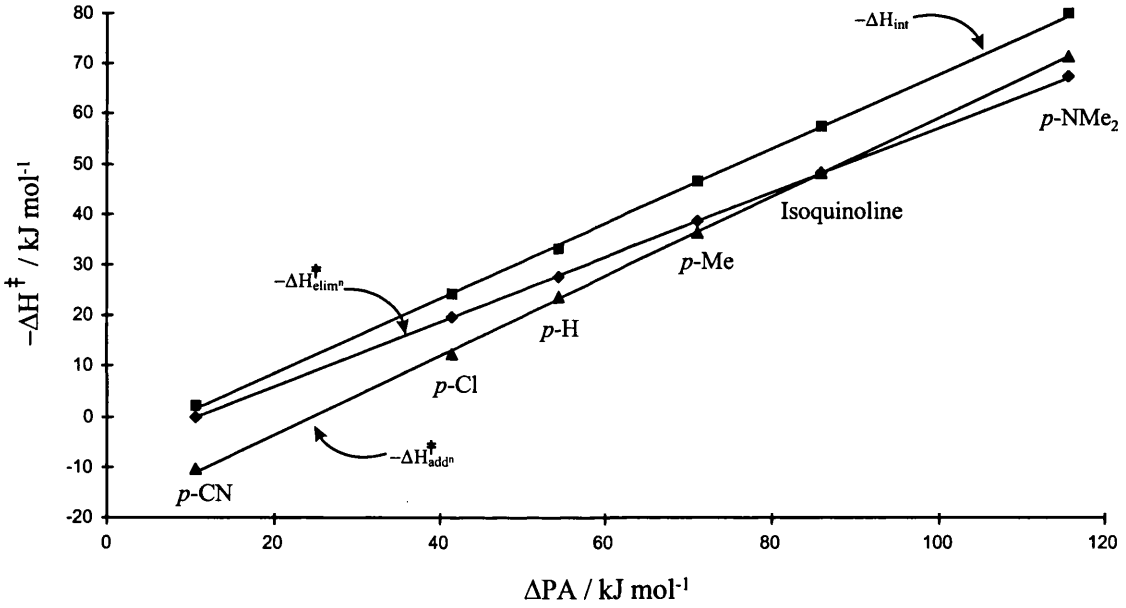


Figure 7.6: Brønsted plot of AM1 heats of activation for the formation and breakdown of the intermediate vs. relative proton affinity of the nucleophile (Hammond & Williams).

Hammond & Williams stated that it was important to discover whether the involvement of an intermediate in the predicted gas-phase mechanism was merely an artefact of the model used or was a true reflection on reality. Their model was deficient in two

respects. First, it neglected the steric bulk of the pyridines and isoquinoline. Secondly, there was no attempt to include the effect of solvation on the system.

7.4 Methods

The transfer of the methoxycarbonyl group between isoquinoline and substituted pyridines in aqueous solution was ideal to be subjected to theoretical exploration, Equation 7.3. Due to the size of the system, the AM1 semi-empirical Hamiltonian in MOPAC93[12] was used to model the 'whole' reacting system. The main deficiency of the earlier theoretical reaction model was that no attempt was made to include the effect of solvation on the system. This is addressed here using the continuum model, COSMO, to mimic the effects of solvation.[11] In the transfer reaction of the methoxycarbonyl group between isoquinoline and pyridine nucleophiles, we are not dealing with substituted aryl compounds so it is appropriate to use a Brønsted rather than the Hammett relationship used in Chapter 5 and 6, i.e. using pK_as instead of σ values.

7.5 Results

The energies for the first transition-state (TS1), the intermediate (INT) and the second transition-state (TS2) are calculated for each member of a series where the substituent X was varied in the nucleophile, Figure 7.2. The enthalpy of activation for the transition-states and intermediate was defined as:

$$\Delta H^{\ddagger}_{(TS)} = \Delta H_{f(TS)} - (\Delta H_{f(Reactant)} + \Delta H_{f(Nucleophile)}) / \text{kJ mol}^{-1} \qquad \text{Equation 7.10}$$

Table 7.1 contains the energies of activation for the transition-states and intermediate for the reaction together with the carbon-nitrogen bond lengths. Table 7.2 contains Pauling bond orders for the reaction where the proportionality constant c is equal to 0.6 together with the charge on the nucleophile, q_{Nu}, defined as the sum of the atomic charges on the nucleophilic group.

Table 7.1: AM1/COSMO calculated energies, kJ mol^{-1}, and bond lengths for the reactions of substituted pyridines with *N*-(methoxycarbonyl) isoquinolinium ion.

X	$\Delta H^{\ddagger}_{TS1}$	$\Delta H^{\ddagger}_{INT}$	$\Delta H^{\ddagger}_{TS2}$	C-$N_{nu_{TS1}}$ / Å	C-$N_{lg_{TS1}}$ / Å	C-$N_{nu_{INT}}$ / Å	C-$N_{lg_{INT}}$ / Å	C-$N_{nu_{TS2}}$ / Å	C-$N_{lg_{TS2}}$ / Å
NH_2	98.5	69.3	75.0	1.868	1.490	1.517	1.584	1.477	1.796
MeO	107.9	91.1	102.1	1.844	1.492	1.538	1.570	1.489	1.822
Me	107.4	91.2	102.6	1.844	1.492	1.542	1.569	1.491	1.826
H	107.7	92.4	103.9	1.831	1.494	1.565	1.565	1.494	1.831
* 	111.5	101.0	111.5	1.843	1.492	1.543	1.567	1.492	1.827
MeOCO	115.6	109.3	121.9	1.809	1.496	1.561	1.557	1.498	1.842
CN	115.6	106.7	121.6	1.812	1.496	1.573	1.548	1.495	1.853
NO_2	123.6	120.0	135.7	1.784	1.499	1.571	1.551	1.502	1.846

* isoquinoline is the attacking nucleophile

Table 7.2: AM1/COSMO calculated Pauling bond orders and the charge on the nucleophile for the TS1, INT and TS2.

X	$n_{C\text{-}N\ TS1}$	$n_{C\text{-}N\ INT}$	$n_{C\text{-}N\ TS2}$	$q_{Nu\ TS1}$	$q_{Nu\ INT}$	$q_{Nu\ TS2}$
NH_2	0.47	0.85	0.91	0.34	0.78	0.86
MeO	0.50	0.84	0.91	0.36	0.76	0.85
Me	0.51	0.84	0.91	0.38	0.77	0.86
H	0.51	0.84	0.91	0.38	0.76	0.86
MeOCO	0.55	0.82	0.92	0.40	0.74	0.87
CN	0.54	0.81	0.92	0.40	0.69	0.87
NO_2	0.57	0.81	0.91	0.43	0.72	0.87

7.6 Discussion

7.6.1 Barrier heights

The examination of the calculated enthalpy of activation for TS1, INT and TS2, shows that for the most powerful nucleophile, X=NH_2, the first transition-state is the rate determining step. As the series is descended to the weakest nucleophile, X=NO_2, the second transition state becomes rate determining, Table 7.1. The reason is that the height of the TS1 barrier is related to the nucleophilicity of the attacking nucleophile. The electron-donating substituents provide more powerful nucleophiles due to their ability to increase the availability of the lone pair for nucleophilic attack. This ability decreases as the substituent becomes more electron-withdrawing. The intermediate is stabilised greatly by electron-donating substituents that can delocalise the positive

charge that is formed on the nitrogen, Figure 7.7. The intermediate is lower in energy than either the TS1 and TS2 barrier heights. For the nucleophilic attack by *p*-nitropyridine, the TS1 barrier height is only 3.6 kJ mol^{-1} higher in energy than its intermediate. By adding more electron-withdrawing groups to the pyridine ring, it may be possible to lower the barrier to such an extent that it could be classed as an 'enforced' concerted mechanism. A further lowering of the barrier would then lead to the reaction being classed as a 'true' concerted mechanism. Table 7.2 shows that as the series is descended, the trend in the bond order, $n_{C\text{-}N}$, and relative charge, q_{Nuc}, of TS1 increases as the bond between the nucleophile and reacting carbon decreases. For TS2, the bond order and relative charge for the nucleophile stays constant as the bond length is close to that which is in the final product.

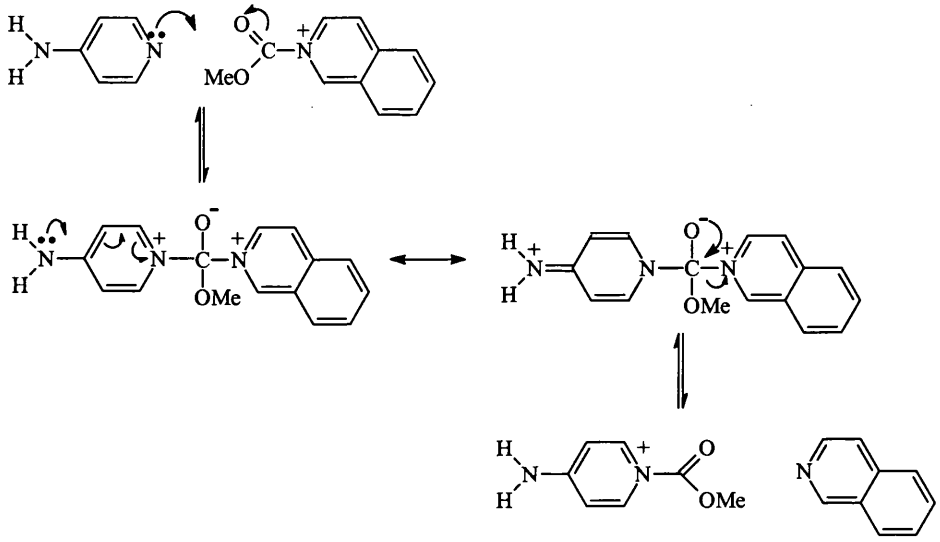
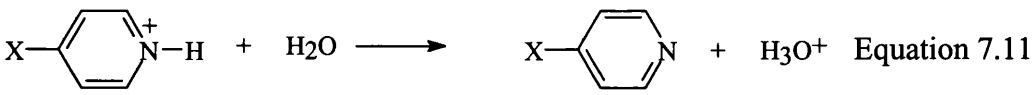


Figure 7.7: The possible transfer mechanism of the methoxycarbonyl group between isoquinoline and *p*-aminopyridine depicting the stabilisation of the intermediate.

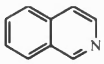
7.6.2 Brønsted correlations



Equation 7.11

The enthalpy for reaction for the proton transfer reaction (PA) where X is a variable substituent, Equation 7.11, are calculated and plotted against $-\Delta H^{\ddagger}$, Table 7.3. A plot of all the barrier heights for the reaction shows a typical stepwise Brønsted correlation, Figure 7.8; The plot shows two intersecting lines with a break at the ΔPA of isoquinoline.

Table 7.3: AM1/COSMO calculated Energies, kJ mol^{-1}, for the reactions of substituted pyridines with *N*-(methoxycarbonyl)isoquinolinium ion.

X	$-\Delta H^{\ddagger}_{TS1}$	$-\Delta H^{\ddagger}_{INT}$	$-\Delta H^{\ddagger}_{TS2}$	$-\Delta H^{\ddagger}_{PRODUCT}$	ΔPA
NH_2	-98.5	-69.3	-75.0	48.1	117.3
Me	-107.4	-91.2	-102.6	10.7	101.2
H	-107.7	-92.4	-103.9	8.6	99.9
	-111.5	-101.0	-111.5	0.0	94.8
MeOCO	-115.6	-109.3	-121.9	-12.9	88.6
CN	-115.6	-106.7	-121.6	-11.3	86.7
NO_2	-123.6	-120.0	-135.7	-27.0	78.5
Correlation	0.99	0.99	1.00	0.99	- -

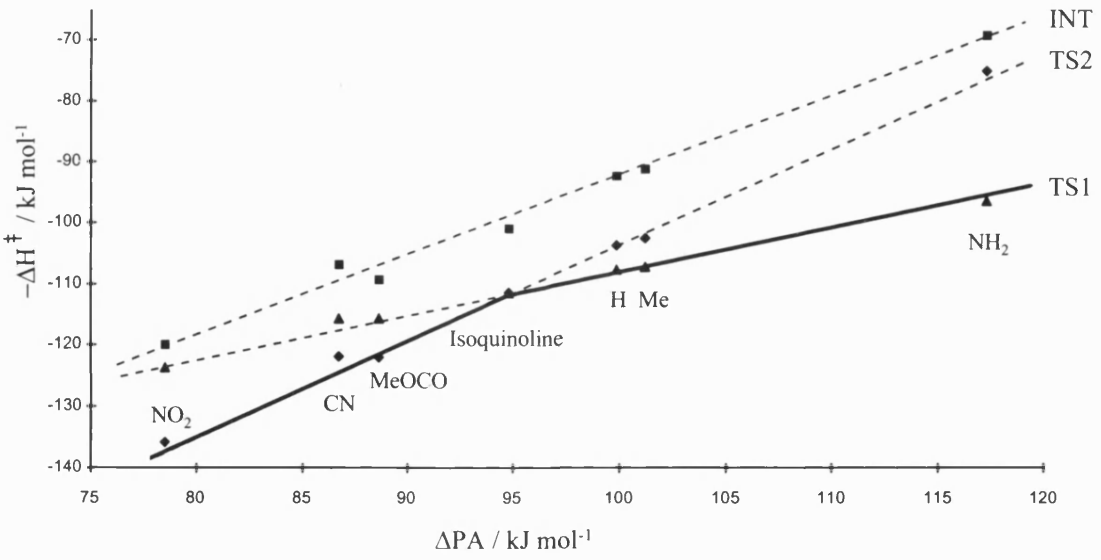


Figure 7.8: A stepwise Brønsted plot of the rate determining step for the reaction of substituted pyridines with *N*-(methoxycarbonyl)isoquinolinium ion.

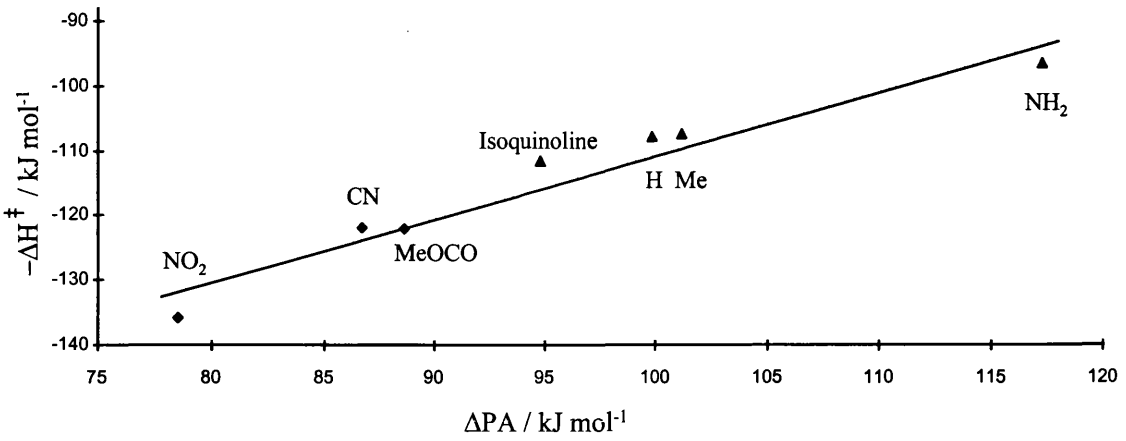


Figure 7.9: A linear Brønsted plot of the rate determining step for the reaction of substituted pyridines with *N*-(methoxycarbonyl)isoquinolinium ion.

A plot of only the rate determining steps, i.e. the larger barrier heights, for each nucleophile gives a Brønsted plot that shows a good linear relationship and correlation (0.96) over a wide pK_a range, Figure 7.9. This could mistakenly be taken as evidence for a concerted reaction mechanism even though the theoretical calculations have clearly shown the reaction to be stepwise. It should be noted that a curve could also be plotted through the data. In the study by Chrystiuk and Williams the large amount of scatter should be noted.

7.7 Conclusion

In the transfer reaction of the methoxycarbonyl group between isoquinoline and substituted pyridines in aqueous solution, Chrystiuk and Williams came to the conclusion that the reaction was concerted and not stepwise in nature due to the fact that their Brønsted plot followed a 'good' linear relationship over a wide pK_a range. In both the AM1 and AM1/COSMO theoretical simulation of the system, the conclusion drawn was that the mechanism was stepwise. A Brønsted plot featuring only the rate determining step gave a reasonably good linear plot which could wrongly be taken as evidence for a concerted mechanism. Great care must be taken when assigning mechanisms as to whether a Brønsted plot is linear or possesses a break. From their results, Chrystiuk and Williams were correct in stating that there is no statistical justification for a stepwise reaction mechanism. However by then inferring that the

reaction must be concerted may be incorrect. It has been shown that by statistical manipulation of the results, incorrect conclusions can be drawn.

The same type of methodology could be used to investigate the nature of change in mechanisms, if any, using the either a phosphyl or sulphyl group transfer reactions for which experimental results are available.[85,86,87] Alternatively, different leaving groups or nucleophiles could be investigated. To test the reliability of the COSMO method, a simulation that includes specific solvent-solute interactions by the utilisation of a hybrid QM/MM method could be used to check the reliability of the COSMO conclusions. A semi-empirical QM/MM transition-state refinement method such as the GRACE/CHARMM[60,15] suite of programs could be used which would give an insight into the important specific stabilisation effect of the solvent.[88,89]

Chapter Eight

The Evaluation of More O'Ferrall- Jencks and Pross-Shaik Models for Bimolecular Substitution Reactions

Chapter 8- The Evaluation of More O'Ferrall-Jencks and Pross-Shaik Models Bimolecular Substitution Reactions

8.1 Introduction

In the study of reaction mechanisms it is helpful to be able to predict the effects of reactant variation on the transition-state structure. The investigation of nucleophilic substitution reactions at benzyl derivatives can be explored using a model based upon the properties of quantum chemically calculated energies for reacting systems. The changes can be rationalised at a descriptive level using two dimensional reaction co-ordinate diagrams and a simple model to explain the shifts in transition-state structures as developed by More O'Ferrall-Jencks.[43,44] This model is attractive because of its relative simplicity and its ability to rationalise and predict experimental data for a wide variety of organic reactions. However, the model is not universally accepted and used. Changes in transition-state structure may also be explained by the valence bond configuration mixing model of Pross and Shaik,[90] in which the reaction co-ordinate is generated by a linear combination of valence bond electronic configurations.

The predictions of the More O'Ferrall-Jencks reaction co-ordinate model and the model proposed by Pross and Shaik will be compared with the observations from the computational study for the bimolecular substitution reaction at benzyl derivatives. These comparisons will allow a determination of the effectiveness of the More O'Ferrall-Jencks and Pross and Shaik theories in the prediction and rationalisation of structure-reactivity effects on organic reactions and will highlight any deficiencies that they possess.

8.1.1 More O'Ferrall-Jencks

A description of a More O'Ferrall-Jencks diagram was provided in Chapter 4. A Moore O'Ferrall-Jencks potential energy surface for an S_N2 reaction mechanism can be seen in Figure 8.1. The reactants are designated as C-X and Nuc^- for the reactant and nucleophile, respectively. The C signifies the α-carbon at the substitution centre of C-X. The reactants are at an energy minimum at the front right corner. At the front left corner is the triple ion pair, $Nuc^-+C^++X^-$. At the back right corner is the hypothetical

intermediate corner, [Nuc–C–X]$^-$, which has a pentacovalent carbon at its centre. The products C–Nuc and X$^-$ are at the back left corner. In the S_N2 mechanism, the C–X bond breaking and C–Nuc bond making process occur at the same time. Therefore the reaction co-ordinate (dark heavy line) runs diagonally from the reactant to the product corner, with the transition-state (TS), which for simplicity is assumed to be at the midpoint.

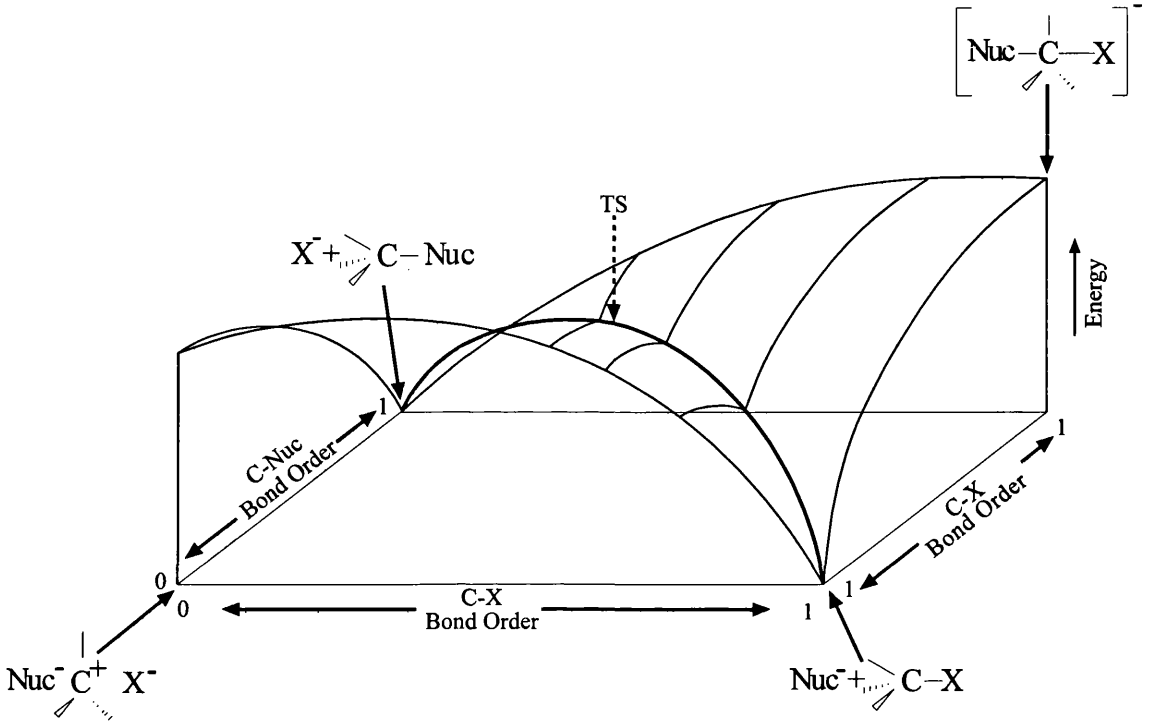


Figure 8.1: 3-Dimensional energy surface and reaction co-ordinate diagram for the S_N2 reaction mechanism.[91]

The effects of various parameters on the location of the S_N2 transition-state can be understood from these diagrams.[64] Changing the substituents so as to supply electrons at the substitution centre will lower the energy of the ion pair corner ($Nuc^-+C^++X^-$). This change will cause a shift of transition-state structure perpendicular to the reaction co-ordinate to TS*, Figure 8.2. The new transition-state will be looser than the first one. The C–X bond cleavage has proceeded farther and the C–Nuc bond formation bond has proceed less than in TS. This would be reflected in greater positive charge on the α-carbon and greater negative charge on the leaving group. Replacing the leaving group by a better one would be equivalent to lowering the energy along the triple ion-reactant edge of the diagram. For the S_N2 process, changing to a better leaving group has two effects on the movement of the transition-state structure. Movement from the centre of

the diagram (TS) back along the reaction co-ordinate towards the reactant corner is accompanied by perpendicular movement of the reaction co-ordinate towards the triple ion corner, ($Nuc^-+C^+X^-$). If both of these motions occur to an equal extent, the net result expected is a new transition-state structure at TS**. The new transition-state will be looser with respect to the C–Nuc distance and approximately unchanged for the C–X distance.

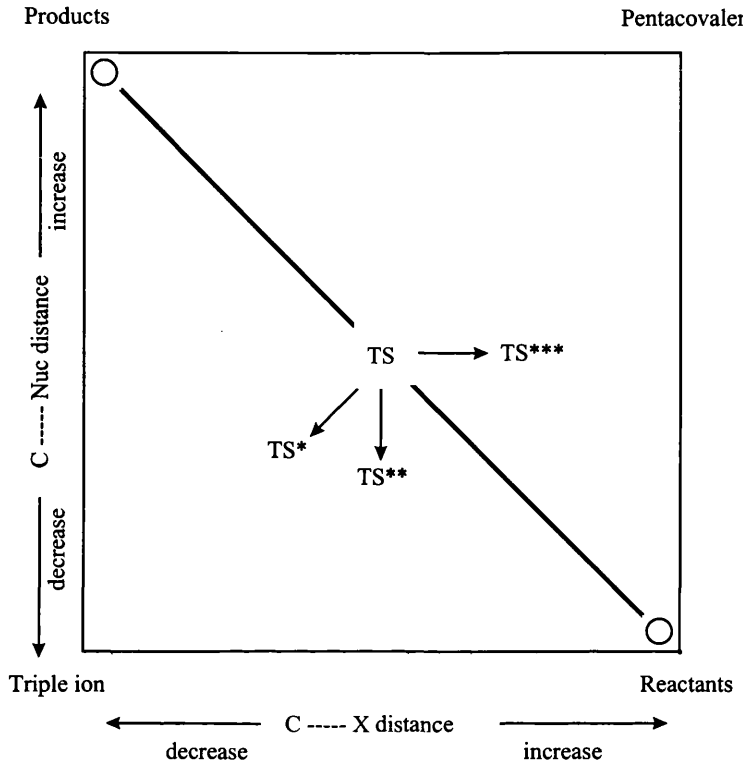


Figure 8.2: A horizontal plane projection of the S_N2 reaction co-ordinate. The transition-state TS is shifted to TS* by an increase in the electron donation to the reacting centre or to TS** by changing to a better the leaving group or to TS*** by changing to a better nucleophile.

The increase in nucleophilicity of the attacking nucleophile means that the reactant-pentacovalent corner will be lowered giving rise to a perpendicular motion while lowering the product corner will give a parallel motion. A new transition-state at TS*** which is the resultant of the perpendicular and parallel vectors. The transition-state will be characterised by a reduction in the R–X cleavage and unchanged with respect to the C–Nuc distance. As a consequence of the reduction in the R–X cleavage there will be a decrease in the charge at both R and X.

8.1.2 Valence bond configuration model

The changes in transition-state structure can be explained by the valence bond configuration mixing model of Pross and Shaik, in which the reaction co-ordinate is generated by a linear combination of valence bond electronic configurations.[92,93]

$$Nuc:^- + R—X \rightarrow Nuc—R + X:^- \qquad \text{Equation 8.1}$$

There are four electrons involved in the S_N2 reaction mechanism, Equation 8.1, and four energetically sensible valence bond forms configuration 8.2a to d, resulting from different distributions of these four electrons. Neither nucleophile (Nuc) or leaving group (X) is given a positive charge as both are more electronegative than R. These are the four important contributors to the reaction profile.

	$Nuc:^- R\cdot \cdot X$	$Nuc\cdot R\cdot :X^-$	$Nuc:^- R^+ :X^-$	$Nuc\cdot R^- \cdot X$
configuration 8.2	a	b	c	d

For a methyl derivative, the S_N2 reaction profile to a first approximation can be built up from configuration 8.2a and b only. This is because the ground state of the reactant can be described by configuration 8.2a and the ground state of the product by configuration 8.2b which puts a full negative charge on either the Nuc or X group. Configurations 8.2c and d place a full positive or negative charge on the CH_3 group and are expected not to have any significant effect on the reaction co-ordinate because they are of such high energy.

At the beginning of the reaction, the nucleophile will be far away from the reactant, configuration 8.3a, and at the end of the reaction, the structure corresponding to same electronic configuration is shown in configuration 8.3b.

	$Nuc:^-$ $R\cdot \cdot X$	$Nuc\cdot \cdot R\cdot^- \cdot X$
configuration 8.3	a	b

It is important to note that according to this model the $Nuc:^-$ structure in configuration 8.2a always has possession of its two electrons no matter if it is close or far from R. There are two reasons why configuration 8.2a increases in energy along the reaction co-ordinate. First, as the distance between R and X increases, the R–X bond breaks leaving the radical X·. Secondly, as the distance between R and Nuc decreases, a three electron repulsive interaction between $Nuc:^-$ and R· is generated. Similarly electron configuration 8.2b decreases in energy from the start of the reaction to the end. In this valence bond structure, $X:^-$ always has possession of two of its electrons. At the

beginning of the reaction, the structure corresponding to this configuration is given by configuration 8.4a at the start and configuration 8.4b at the end. Configuration 8.4b does not have the three-electron repulsion that exists in configuration 8.4a as it has a bond between Nuc and R.

configuration 8.4

Nuc· ·R· ·X⁻ (a) Nuc· ·R :X⁻ (b)

The changes in energy of configuration 8.2a and b along the reaction co-ordinate for a thermoneutral and exothermic reaction can be seen in Figure 8.3. The dashed line in each case shows the reaction profile for the reaction. The reaction profile follows the potential curve for configuration 8.2a before the transition-state and configuration 8.2b after the transition-state. The transition-state is the 'avoided-crossing' of the two configurations.

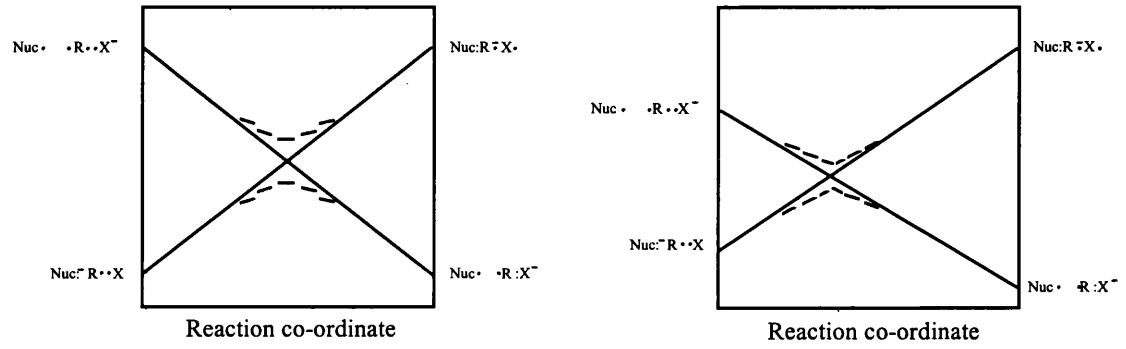


Figure 8.3: A simplified S_N2 reaction profile from configuration 8.2a reactant to configuration 8.2b product. The dashed lines denote the avoided crossing for (a) a thermoneutral and (b) an exothermic reaction.

At the transition-state of Equation 8.1 there is a single electron shift from Nuc:⁻ to X. This is true since configuration 8.2a Nuc:⁻ has two electrons and X· has one but in configuration 8.2a Nuc· has one and X:⁻ has two. The transition-state is the avoided crossing of the potential energy curves of configuration 8.2a and b. This is where the reacting species can be considered to switch over from configuration 8.2a to b. According to this model there is never any charge on R. R has one electron with which it forms a bond with X in configuration 8.2a and with which it forms a bond with Nuc in configuration 8.2b. Nuc:⁻ has a whole negative charge early on in the reaction while X:⁻ has a whole negative charge later on in the reaction. At the transition-state X and Nuc, each have half a negative charge. In the exothermic reaction of Figure 8.3b the

transition-state, in agreement with the Hammond postulate, occurs earlier in the reaction co-ordinate.

In Figure 8.3 the energies of configuration 8.2a and b change linearly along the reaction co-ordinate. This is an over-simplification and may not be the case. At the start of the reaction, when the nucleophile is far from RX and X is close to R, the first excited state of the reactants, configuration 8.2b, is bonded as shown in configuration 8.4a. In the excited state the nucleophile is a free radical and the extra electron on R is in a three electron bond with X. At the end of the reaction the first excited state of the products, configuration 8.2a, is bonded as shown in configuration 8.3b. Now the free radical and R–Nuc is a three electron bond. If the three electron bonds are stable, which means that they are extensively delocalised, the curves may descend less steeply at the edges of the diagram than in centre. This situation is shown in Figure 8.4(a). On the other hand if they particularly unstable the curves may descend more steeply to the avoided crossing as shown Figure 8.4(b).

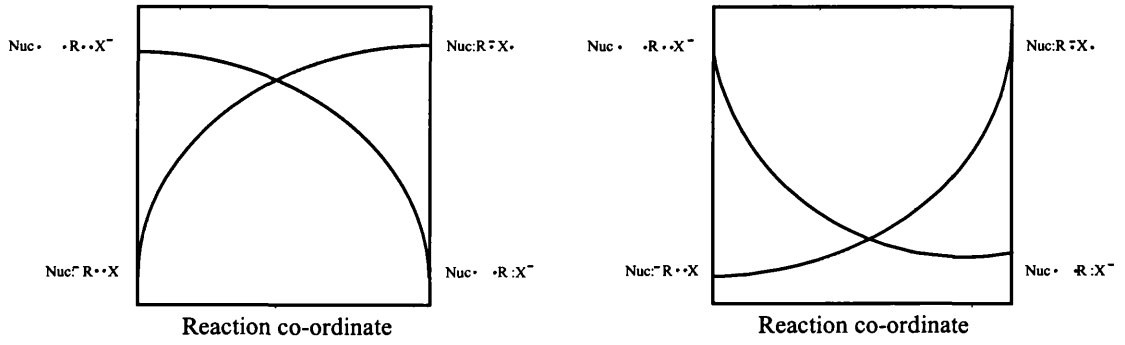
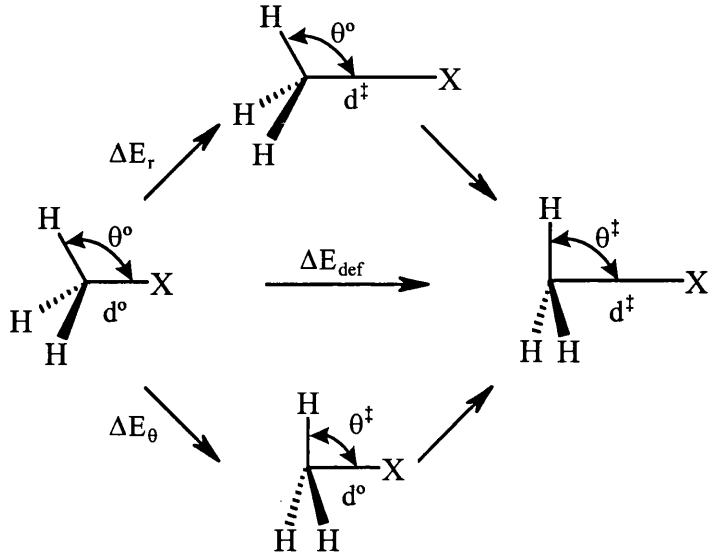


Figure 8.4: (a) An S_N2 profile for the case when the first excited states of the reactant and product are stabilised. (b) An S_N2 profile for the case when the first excited states of the reactant and product are destabilised.

A chemical reaction involves both geometric and electronic reorganisation. The barrier to a chemical reaction comes about from the way these two reorganisations are coupled together. The barrier arises from the need to distort the reactant structure in such a way that simultaneously lowers the energy of the product configuration and raises the energy of the reactant configuration to a point at which they are isoenergetic. This geometric distortion is the mechanism by which the electronic rearrangement can take place.[94] The overall deformation energy required for the reactant to reach its transition-state can be partitioned into ΔE_r which is associated with the bond stretch, and into a component

ΔE_θ which is associated with angular deformation, as shown in Equation 8.5. From their calculations the main component of the deformation energy is the ΔE_r term which is related to the C-X stretch. It was proposed that there should therefore be possible to link, the looseness of a transition-state structure to the magnitude of the reaction barrier.[95]



Equation 8.5

The distortion which distinguishes the reactant and the transition-state from each other is defined as $\Delta d^{\ddagger} = d^{\ddagger} - d^{\circ}$. To create a uniform scale of relative C-X stretching for different Xs, it was convenient to define the percentage of C-X stretching at the transition-state, Equation 8.6.

$$\%CX^{\ddagger} = 100\ \Delta d^{\ddagger} / d^{\circ} \qquad \text{Equation 8.6}$$

The looseness of a system can be defined as the sum of the percentages for the two C-X bonds, as shown in Equation 8.7.

$$\%L^{\ddagger} = 2\ (\%CX^{\ddagger}) \qquad \text{Equation 8.7}$$

A large $\%L^{\ddagger}$ is characteristic of a loose or 'exploded' transition-state. A 'fast' reaction (i.e. small intrinsic barrier) will therefore only require slight distortion to reach its transition-state which will be tight. This model maintains that a large distortion energy corresponds to a looser or "exploded" transition-state. Therefore reactions with a high intrinsic barrier will require a large amount of molecular distortion to reach the transition-state which therefore will be loose.[96]

The height of the transition-state depends on four factors. Other factors being equal, Figure 8.3 shows that the barrier heights is related to the exothermicity of a reaction.[97]

The second factor is the degree to which the crossing is avoided. At the point of crossing, the two states are of equal energy and therefore they mix into each other strongly, causing the lower energy one to be lowered still further and the higher energy one to be raised further. Thus the transition-state is at a lower energy than the crossing of the two configurations and this difference is termed β. The third factor is the height of the energy barrier between the ground state and excited state curves at the start of the reaction co-ordinate. The energy difference in going from the lower to the upper curve, is due the fact that an electron must be removed from Nuc and added to R–X. Thus the difference in energy between the ground state and excited state depends on $I_{Nuc}-A_{RX}$ in which I_{Nuc} is the ionisation potential of the nucleophile and A_{RX} is the electron affinity of R–X. The strength of the three electron bond of the first excited state of the reactants can affect the reactivity of the system. As Figure 8.4 shows, that for a given $I_{Nuc}-A_{RX}$ value, if the bond is strong and delocalised, the avoiding crossing will be reached at a higher energy than if the bond is weak. The factors that effect the energy of the transition-state in S_N2 reactions can be summed up as $E = f(I_{Nuc}-A_{RX})-\beta$. The factor f is the fraction of the energy gap $I_{Nuc}-A_{RX}$ that is under the avoided crossing. The fraction depends on both the shapes of the first excited state curves and on ΔH of the reaction.[98]

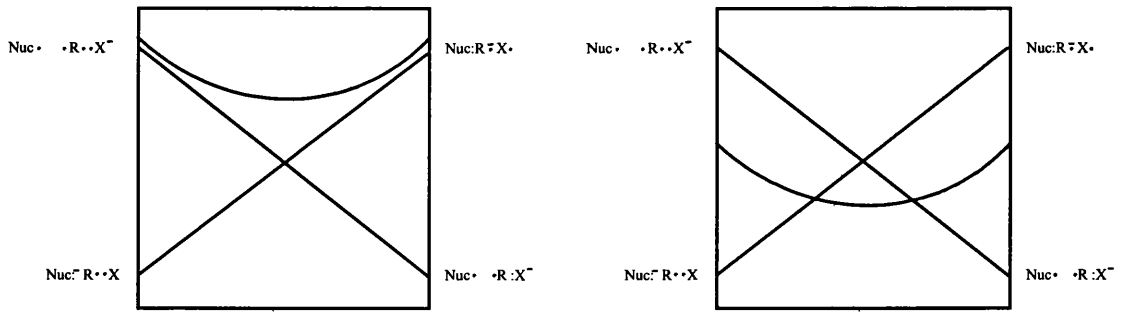


Figure 8.5: (a) Reactant, product and carbocation configurations involving a moderately stable R^+. (b) Reactant, product and carbocation configurations where R^+ is stable enough that an intermediate is formed.

To discuss the S_N2 reactions on benzyl derivatives by the Pross and Shaik method, we need to build the reaction profile from the original valence bond structures, configuration 8.2a to d. However in this case we can not discard configuration 8.2c and d as we did for the methyl derivative. For compounds containing an electron donating

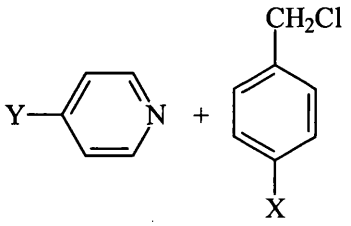
substituent, configuration 8.2c will be important. We can build reaction profiles as seen in Figure 8.5 by plotting the energies of configuration 8.2a to d. Configuration 8.2a and b will behave as in the S_N2 reaction of methyl derivatives. Configuration 8.2c is high in energy at the start of the reaction but halfway along the reaction co-ordinate when $Nuc:^-$ draws close to R^+ and R^+ is now surrounded by two anions, it will possess a potential energy minimum. As X^- departs, the potential energy rises again. Depending on the stability of configuration 8.2c at its potential minimum, the potential energies of configuration 8.2a to c may vary as shown in Figure 8.5a or the reaction will have a low barrier and involve a reaction intermediate whose electronic configuration corresponds to configuration 8.2c. When configuration 8.2c is mixed in with configuration 8.2a and b, they no longer contribute equally to the transition-state. In Figure 8.5b, the transition-state occurs when configuration 8.2a and c are of equal energy but configuration 8.2b is still of much higher energy and thus is only a small contributor.

The effects of various parameters on the location of the S_N2 transition-state can be understood by examining the effects they have on configurational stability. The result of increasing the leaving group ability is to lower the energy of configuration 8.2b and c in which the R–X bond pair is already delocalised on the leaving group. The stabilisation of configuration 8.2b and c leads to an increased bond loosening of the R–X bond since both configurations favour a weak R–X linkage. The effect on the Nuc–R bond is less clear as configuration 8.2b favours bond formation whereas configuration 8.2c favours little Nuc–R bond making. Configuration 8.2b mixes into the transition-state to a greater extend than configuration 8.2c due to its greater stability of the former configuration. As a consequence the effect of introducing a better leaving group is expected to lead to a greater Nuc–R bond formation. The result of increasing the nucleophile strength is the lowering of configuration 8.2b since the nucleophile is a better donor consequently mixing into the transition-state to a greater extent. The configuration 8.2b encourages greater Nuc–R bond formation as well as greater R–X bond cleavage.

The effect of ring substituents in the ring will have a significant impact on the relative stabilities of the different configurations. Configuration 8.2a has a slight positive charge

on the carbon which is a reflection of the polar R–X bond. Configurations 8.2b and d possess, a slight negative charge on the carbon while configuration 8.2c has a positive charge on the carbon. Thus the effect of electron donating groups will be to stabilise configuration 8.2c at the expense of the other configurations. The increased participation of configuration 8.2c will lead to a looser transition-state for both the Nuc–R and R–X bonds, leading to a transition-state with greater positive charge on the α-carbon. As configuration 8.2a and b are raised in energy by electron donating groups, the Nuc–R and R–X bonds will be weaker and therefore the transition-state will be looser.

8.1.3 Method



X = NH_2, H or NO_2

Y = NH_2, MeO, Me, H, Br, CN or NO_2

Figure 8.6: The bimolecular nucleophilic reaction with variable substituted benzyl chloride and pyridine nucleophiles.

The reaction chosen for this study is the bimolecular nucleophilic substitution reaction of a family of *p*-amino, unsubstituted and *p*-nitro benzyl chlorides with varying pyridine nucleophiles, Figure 8.6. The various pyridine nucleophiles are chosen so that the system will be subjected to nucleophiles with variable nucleophilicity and therefore the β_{nuc} for the different families can be calculated. The AM1 semi-empirical Hamiltonian utilising COSMO,[11] a continuum model, in MOPAC93,[12] is used to model the reaction in water. The CAMVIB and CAMISO programs are used to compute the kinetic isotope effects.[28]

8.2 Results

The activation energy for the bimolecular nucleophilic substitution reactions are calculated for the series of different substituted benzyl chlorides with varying pyridine nucleophile, Table 8.1. The activation energy for the bimolecular transition-state is defined as:

$$\Delta H^{\ddagger}_{(S_N2)} = \Delta H_{f(S_N2\ TS)} - \{\Delta H_{f(Y\text{-}Py)} + \Delta H_{f(X\text{-}BzCl)}\} \ / \ kJ\ mol^{-1} \qquad \text{Equation 8.8}$$

The transition-state bond lengths (Å) for the carbon-chlorine (C–Cl) and carbon-nitrogen (C–N) distance are given together with the Pauling bond orders which are calculated in the same manner as described in Chapter 4. The calculated charge on the nucleophile, leaving and benzyl groups at the transition-state are given in Table 8.2. The charge on the benzyl group is defined as the total charge on the benzyl group plus that of the substituents. Table 8.3 contains the α-deuterium, carbon, chlorine leaving group and nitrogen nucleophile kinetic isotope effects at 298.15K. The α-deuterium kinetic isotope effects are given per deuterium atom.

Table 8.1: AM1/COSMO calculated activation energies, kJ mol^{-1}, transition-state bond lengths and Pauling bond orders.

Substituent X	Substituent Y	$\Delta H^{\ddagger}_{S_N2}$	C–Cl / Å	C–Nuc / Å	$n_{C\text{–}Cl}$	$n_{C\text{–}Nuc}$
NH_2	NH_2	139.6	2.90	2.07	0.15	0.36
NH_2	MeO	140.3	2.91	2.06	0.15	0.37
NH_2	Me	140.6	2.91	2.06	0.15	0.37
NH_2	H	141.7	2.91	2.06	0.15	0.38
NH_2	Br	143.6	2.92	2.04	0.15	0.39
NH_2	CN	145.0	2.93	2.03	0.15	0.40
NH_2	NO_2	147.3	2.94	2.02	0.15	0.40
H	NH_2	161.7	2.39	2.05	0.36	0.38
H	MeO	162.9	2.40	2.04	0.35	0.39
H	Me	163.4	2.40	2.04	0.36	0.39
H	H	164.2	2.41	2.04	0.35	0.39
H	Br	166.5	2.42	2.03	0.35	0.40
H	CN	169.7	2.42	2.02	0.34	0.40
H	NO_2	171.0	2.43	2.01	0.34	0.41
NO_2	NH_2	171.1	2.30	2.01	0.42	0.40
NO_2	MeO	172.9	2.31	2.00	0.42	0.41
NO_2	Me	173.2	2.31	2.00	0.42	0.41
NO_2	H	173.9	2.31	2.00	0.42	0.41
NO_2	Br	177.0	2.32	1.99	0.41	0.42
NO_2	CN	180.5	2.32	1.98	0.41	0.43
NO_2	NO_2	182.3	2.33	1.97	0.40	0.44

Table 8.2: AM1/COSMO charge distribution in the transition-state for the nucleophile, leaving and benzyl group.

Substituent X	Substituent Y	Nucleophile	Leaving Group	Benzyl Group
NH_2	NH_2	0.16	-0.99	0.83
NH_2	MeO	0.16	-0.99	0.83
NH_2	Me	0.16	-0.99	0.83
NH_2	H	0.16	-0.99	0.83
NH_2	Br	0.18	-0.99	0.81
NH_2	CN	0.18	-0.99	0.80
NH_2	NO_2	0.19	-0.99	0.80
H	NH_2	0.17	-0.86	0.69
H	MeO	0.19	-0.87	0.68
H	Me	0.19	-0.87	0.68
H	H	0.19	-0.87	0.68
H	Br	0.20	-0.88	0.68
H	CN	0.20	-0.88	0.68
H	NO_2	0.20	-0.88	0.68
NO_2	NH_2	0.17	-0.81	0.64
NO_2	MeO	0.18	-0.81	0.63
NO_2	Me	0.19	-0.81	0.62
NO_2	H	0.20	-0.81	0.61
NO_2	Br	0.20	-0.81	0.61
NO_2	CN	0.20	-0.80	0.60
NO_2	NO_2	0.20	-0.80	0.60

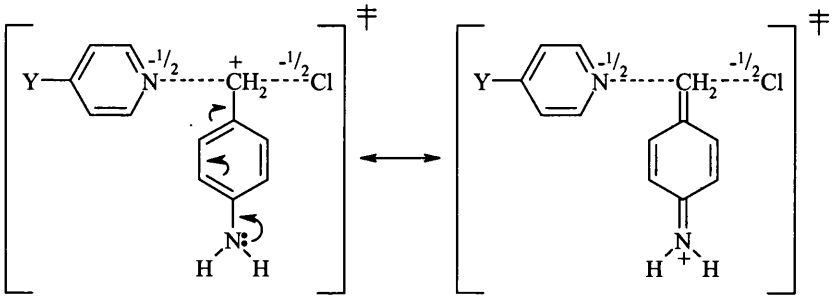
Table 8.3: AM1/COSMO calculated α-deuterium, carbon, chlorine leaving group and nitrogen nucleophile kinetic isotope effects at 298.15 K.

Substituent X	Substituent Y	k_H/k_D	$k_{C^{12}}/k_{C^{14}}$	$k_{Cl^{35}}/k_{Cl^{37}}$	$k_{N^{14}}/k_{N^{15}}$
NH_2	NH_2	1.035	0.9990	1.0147	0.7998
NH_2	MeO	1.035	0.9974	1.0097	0.8030
NH_2	Me	1.037	0.9974	1.0118	0.8027
NH_2	H	1.037	0.9980	1.0092	0.8022
NH_2	Br	1.038	0.9981	1.0091	0.8025
NH_2	CN	1.037	0.9965	1.0090	0.7980
NH_2	NO_2	1.036	0.9935	1.0088	0.7626
H	NH_2	1.009	0.9497	1.0076	0.8031
H	MeO	1.010	0.9499	1.0086	0.8020
H	Me	1.008	0.9441	1.0079	0.8036
H	H	1.011	0.9510	1.0078	0.8013
H	Br	1.009	0.9551	1.0093	0.8046
H	CN	1.011	0.9560	1.0090	0.8025
H	NO_2	1.014	0.9564	1.0084	0.8022
NO_2	NH_2	1.007	1.0028	1.0082	0.8036
NO_2	MeO	1.009	1.0017	1.0069	0.8014
NO_2	Me	1.006	1.0042	1.0102	0.8020
NO_2	H	1.006	1.0023	1.0070	0.8007
NO_2	Br	1.009	1.0081	1.0083	0.8040
NO_2	CN	1.009	1.0094	1.0080	0.7901
NO_2	NO_2	1.009	1.0097	1.0076	0.7618

8.3 Discussion

8.3.1 Variation of the *p*-substituent

The *p*-amino substituent lowers the bimolecular activation energy by approximately 20 kJ mol^{-1} relative to the unsubstituted benzyl chloride, while the effect of the *p*-nitro substituent is to raise it by approximately 10 kJ mol^{-1}. The *p*-amino substituent is able to stabilise the carbocation character of the transition-state by electron resonance donation, Equation 8.9, which removes the positive charge away from the α-carbon, whereas the electron withdrawing *p*-nitro substituent destabilises the transition-state by pulling electron density away from the reacting α-carbon centre.



Equation 8.9

The *p*-amino stabilisation of the carbocation reacting centre causes the carbon-chlorine leaving group and carbon-nitrogen nucleophile bonds to become elongated as stabilisation by the nucleophile and leaving group is less required. Also, the removal of the positive charge from the α-carbon in the second resonance structure increases the repulsion between the ½ negative charges. This gives rise to a transition-state which is loose. The effect is largest for the carbon-chlorine bond which is reflected in the *p*-amino families possessing the greatest α-deuterium isotope effect due to their transition-state looseness. The *p*-nitro destabilisation of the carbocation reacting centre causes the carbon-chlorine and carbon-nitrogen bonds to become truncated as stabilisation by the nucleophile and leaving group is required. This gives rise to a tight transition-state which is reflected in the *p*-nitro families small α-deuterium isotope effects that are close to unity.

Since a bond to carbon is always broken in S_N2 mechanisms, a carbon isotope effect should be observed. Qualitatively a maximum in carbon isotope effect is observed for a transition-state in which the nucleophile and leaving group are both bonded to carbon with equal strength so the tighter the transition-state the greater the value.[22] The

calculated carbon isotope effects show that the largest value occurs for the *p*-nitro family which have the tightest transition-states. It would then be predicted that the *p*-amino family would possess the smallest values as their transition-states are the loosest. However this is not the case as the unsubstituted benzyl chloride family possess the smallest values. The *p*-amino benzyl family possess the greatest positive benzyl charge as it is able to stabilise the charge appropriately whereas the charge for *p*-nitro benzyl family is smaller as it can not stabilise a large positive charge.

8.3.2 Constant chlorine leaving group

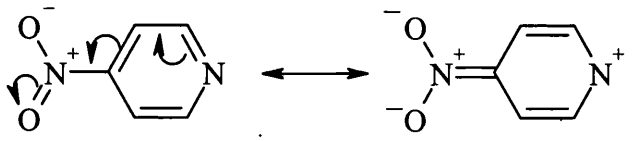
The carbon-chlorine bond length shows little variance within a *p*-substituent benzyl family. The comparison between the different families shows that the *p*-amino family possess the loosest carbon-chlorine bond which becomes tighter as you go through to the *p*-nitro substituent. The charge at the transition-state is a direct reflection of the tightness or looseness of the system. The negative charge on the chlorine leaving group is near unity, a value close to a 'free' chloride anion, for the 'loose' *p*-amino benzyl family. As the carbon-chlorine bond length decreases and the transition-state tightens, going from *p*-amino to the unsubstituted to the *p*-nitro family, there is a decrease in charge for the chlorine leaving group (i.e. it becomes less negative).

The reactions with strong electron donating groups are able to stabilise the carbocation centre, giving rise to more carbon-chlorine bond rupture, and hence larger isotope effects. The calculated chlorine isotope effects show a maximum value of $k_{Cl^{35}}/k_{Cl^{37}}$ = 1.0147 for the reaction between *p*-amino benzyl chloride and p-amino pyridine which possess the loosest carbon-chlorine bond length. Whereas the reaction between *p*-nitro benzyl chloride and *p*-nitro pyridine which possess the tightest carbon-chlorine bond length shows a minimum value of $k_{Cl^{35}}/k_{Cl^{37}}$ = 1.0076. The chlorine leaving group isotope effect, shows only slight overall variance between its maximum and minimum values, indicating that very accurate results are required for the chlorine isotope effects to be a good measurement of transition-state structure.

8.3.3 Variation of the nucleophile

A nucleophilic species has to be capable of sharing an unpaired electron with the atom being attacked. So the greater this ability, the stronger the nucleophile. The nucleophilic ability of a species is of great importance in S_N2 reactions as it is involved in the rate determining step. The nucleophiles used in this study are a family of pyridines whose *p*-substituents have been made either electron donating or withdrawing. If the *p*-substituent is electron donating then it increases the nucleophilicity of the attacking pyridine nucleophile by donating electron density to the attacking nitrogen atom, Equation 8.10. If the *p*-substituent is made electron withdrawing then it decreases the nucleophilicity of the pyridine nucleophile by removing electron density away from the attacking nitrogen atom, Equation 8.11.

Equation 8.10



Equation 8.11

The stronger nucleophiles possess electron donating substituents at the *p*-substituent such as the amino group. This manifests itself in lower activation energies and subsequent looser carbon-nucleophile bond lengths at the transition-state. The weaker nucleophiles possess electron withdrawing substituents at the *p*-substituent such as the nitro group. This manifests itself in higher activation energies and subsequent tighter carbon-nucleophile bond lengths at the transition-state. The nucleophilic charge at the transition-state is positive for all the nucleophiles with the electron donating substituents possessing the smallest positive charge. The *p*-benzyl substituents have little or no effect on the charge of nucleophile at the transition-state. All the nitrogen kinetic isotope effects for the substituted nucleophiles are inverse and become more inverse as the transition-state tightens.

8.4 Comparison and criticisms of the transition-state models

A summary of the findings of this theoretical investigation and of the transition-state models is given in Table 8.4. The MOFJ and Pross and Shaik models correctly predict the changes in transition-state structure by increasing the electron donation to the reacting centre. The transition-state will become looser with respect to the carbon-leaving and carbon-nucleophile bond which is reflected in the greater positive charge on the benzyl group and greater negative charge on the leaving group. The MOFJ theory correctly predicts that the increased nucleophilicity of the attacking nucleophile will lead to a tighter carbon-leaving group bond whereas the Pross and Shaik model incorrectly predicts its loosening.

Table 8.4: A summary of the predicted effects of changing substitutional parameters of an S_N2 reaction mechanism by the More O'Ferrall-Jencks and the valence bond configuration mixing model of Pross and Shaik and their comparison with the calculated results.

Parameters	More O'Ferrall-Jencks	Pross and Shaik	Calculated
Increased leaving group ability	Looser C–Nuc No effect C–X	Tighter C–Nuc Looser C–X	N/A
Increased nucleophilicity	Looser C–Nuc Tighter C–X	Tighter C–Nuc Looser C–X	Looser C–Nuc Tighter C–X
Electron donating substituent[a]	Looser C–Nuc Looser C–X	Looser C–Nuc Looser C–X	Looser C–Nuc Looser C–X
Electron withdrawing substituent[b]	Tighter C–Nuc Tighter C–X	Looser C–Nuc Looser C–X	Tighter C–Nuc Tighter C–X

N/A = not applicable as leaving group was not varied. a = relative to benzyl chloride and an increase in rate, b = relative to benzyl chloride and a decrease in rate, C–X is the carbon-nucleophile bond and C–X is the carbon-leaving group bond.

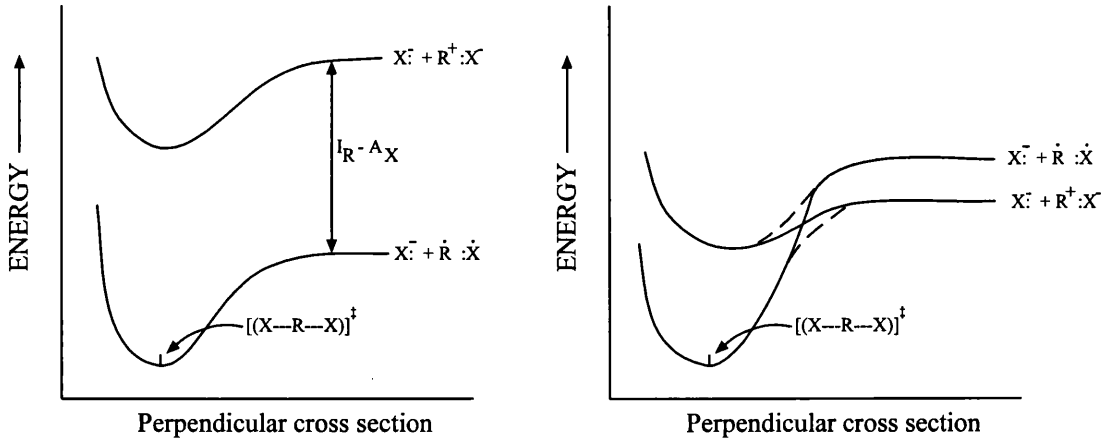
In the MOFJ model, if the parallel and perpendicular motions are equal, then the carbon-leaving bond would tighten and the carbon-nucleophile would be unaffected. However, from the asymmetric and symmetric nucleophile-carbon-chlorine stretches, the motion in the parallel direction is greater than in the perpendicular direction, Table 8.5. Therefore, the resultant vector would give rise to a tightening of the carbon-leaving group bond. The effect of destabilising the benzyl carbocation by electron withdrawing ring substituents is to increase the bonding of the benzylic carbon to both the

nucleophile and leaving group which decreases the positive charge at the benzylic carbocation. The MOFJ model correctly predicts the tightening of the transition-state whereas the Pross and Shaik model incorrectly predicts transition-state loosening.

Table 8.5: Selected asymmetric and symmetric nucleophile-carbon-chlorine bond stretches, cm^{-1}.

Reactant	Nucleophile	asymmetric	symmetric
p-amino benzyl chloride	pyridine	-445.8	976.2
benzyl chloride	pyridine	-589.0	974.5
p-nitro benzyl chloride	pyridine	-691.5	973.0

In a polar reaction, the dissociative corner in a MOFJ diagram refers to a highly ionic species. In nucleophilic reactions, this corner is the triple ion ($Nuc:^- R^+ :X^-$) and in proton transfers the triple ion ($B:^- H^+ :A^-$). In the gas phase, such structures are 300-1250 kJ mol^{-1} higher in energy than radical structures such as ($Nuc:^- R\cdot \cdot X$) and ($B:^- H\cdot \cdot A$) due to the high energy cost of ionisation in the gas phase. A perpendicular cross section from the associative to dissociative corner of a MOFJ diagram for the identity reaction $X^- + CH_3X \rightarrow XCH_3 + X^-$ can be seen in Figure 8.7a. The transition-state may correlate with either the triple ion ($Nuc:^- CH_3^+ :X^-$) or with the ($Nuc:^- CH_3\cdot \cdot X$) fragments. The gas phase triple ion will lie higher in energy by a gap whose value is ($I_R - A_X$), i.e. the ionisation potential of $CH_3\cdot$ minus the electron affinity of $X\cdot$. This gap is typically 300-800 kJ mol^{-1} depending on X. Pross and Shaik argued that in the gas phase, the identity of the dissociative corner is said to be ($Nuc:^- CH_3\cdot \cdot X$) (and its mirror image) and not a triple ion.



a b

Figure 8.7: a) In the gas phase, the dissociative corner of an identity S_N2 reaction showing that it can not be the triple ion species. b) The effect of solvent on the nature of the dissociative corner. The dashed lines represents the avoided crossing.

In solution the triple ion will be more strongly solvated than the (Nuc:$^-$ CH_3· ·X) fragments. This 'solvent assisted' crossing can be seen in Figure 8.7b. However, the curvature of the surface along the perpendicular mode will continue to be influenced by the neutral fragments, and changes in the energy of the carbocation may have unpredictable effects upon the energy of the transition-state. According to Pross and Shaik, the role of the looseness of the dissociative corner is no longer unambiguous. Since the fragments may be either negatively, neutral or positively charged, there is no straightforward relationship between the charge at the transition-state and its looseness. This conclusion conflicts with the calculated results that show that the looser the transition-state is, the greater the positive charge on the benzyl system and the greater negative charge on the leaving group.

The MOFJ diagrammatic model and the qualitative valence bond curve crossing model of Pross and Shaik each predict that the transition-state for nucleophilic substitution reactions at benzyl derivatives will change as the benzyl carbocation is destabilised by electron withdrawing ring substituents. However, the changes predicted by the two models differ. The MOFJ treatment predicts a shift away from the triple ion corner towards the centre of the diagram. According to the conventional view, this corresponds to a decrease in the positive charge at the benzylic carbon and by a tightening of the transition-state with an increase in the bonding of the benzylic carbon to both the

nucleophile and leaving group. This would lead to an increase in the value of β_{nuc} for the reaction.

The Pross and Shaik model predicts that destabilisation of the benzyl carbocation decreases the contribution of the valence bond structure ($Nuc:^{-}R^{+}:X^{-}$) to the transition-state with a consequent decrease in positive charge at the benzylic carbon. However, this model maintains that slow reactions (i.e. those with high intrinsic barriers) require a large amount of molecular distortion to reach the transition-state, which is therefore loose. Thus a change to a more electron withdrawing substituent which would cause a decrease in the rate of a nucleophilic substitution reaction would result in a looser transition-state. This would lead to a decrease in the value of β_{nuc} for the reaction.

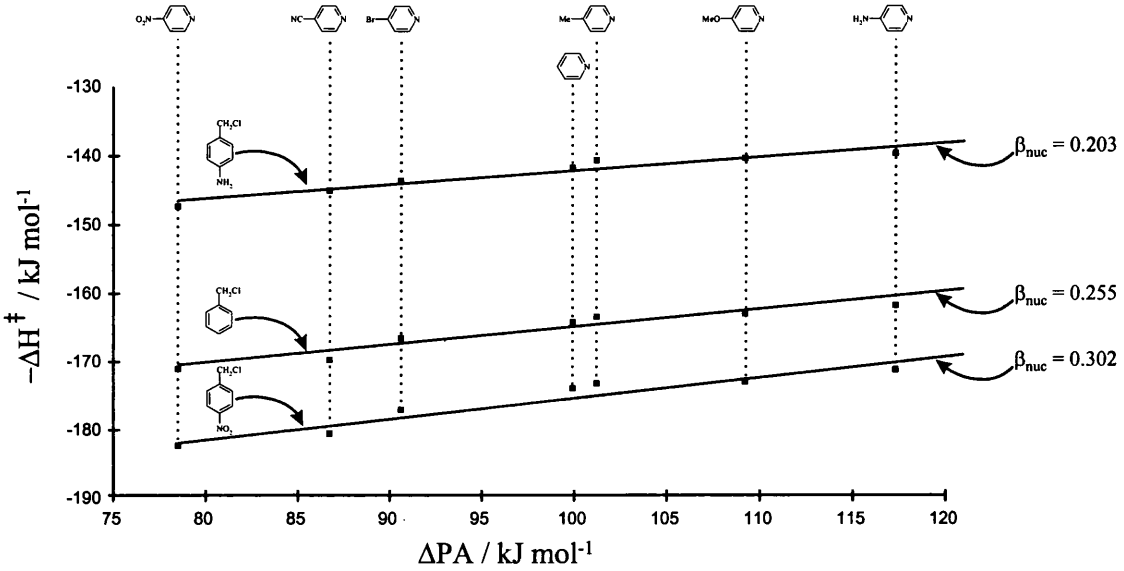


Figure 8.8: Brønsted plot of AM1/COSMO (negative) heats of activation vs. relative proton affinity of the nucleophile.

The calculated results show that the effect of destabilising the benzyl carbocation by electron withdrawing ring substituents is to increase the bonding of the benzylic carbon to both the nucleophile and leaving group which decreases the positive charge at the benzylic carbon. To determine the value of β_{nuc} the negative activation enthalpies vs. relative proton affinity are plotted, Figure 8.8. The β_{nuc} values for the reaction show an increase as the benzyl carbocation is destabilised by the addition of election withdrawing groups. The calculated β_{nuc} values conform to the predicted MOFJ diagrammatic model and not to the valence bond curve crossing model of Pross and Shaik.

The looseness of the leaving group for a non-identity reaction is defined as the percentage of elongation of X at the transition-state relative to the reactant bond length, Equation 8.12. The looseness of the nucleophile is defined as the percentage of elongation of Y at the transition-state relative to the product bond length, Equation 8.13.

$$\%CX^{\ddagger} = 100\% \left[(d^{\ddagger}_{CX} - d^{o}_{CX} / d^{o}_{CX}\right] \qquad \text{Equation 8.12}$$

$$\%CY^{\ddagger} = 100\% \left[(d^{\ddagger}_{CY} - d^{o}_{CY} / d^{o}_{CY}\right] \qquad \text{Equation 8.13}$$

The looseness of a system can then be defined as the sum of the percentages for the two bonds, as shown in Equation 8.14.

$$\%L^{\ddagger} = \%CX^{\ddagger} + \%CY^{\ddagger} \qquad \text{Equation 8.14}$$

Table 8.6: The looseness of transition-states, $\%L^{\ddagger}$, as defined by Pross and Shaik

Nucleophile Substituent Y	*p*-NH_2 family $\%L^{\ddagger}$	*p*-H family $\%L^{\ddagger}$	*p*-NO_2 family $\%L^{\ddagger}$
NH_2	104.8	74.2	66.8
MeO	103.8	73.6	66.0
Me	103.6	73.5	65.9
H	103.5	73.9	66.0
Br	102.5	73.5	65.5
CN	102.3	73.2	64.8
NO_2	101.8	73.1	64.4

The Pross and Shaik definition of looseness are calculated for the studied reaction, Table 8.6. Figure 8.9 is a plot of $\%L^{\ddagger}$, vs. activation energy for the AM1/COSMO data. The plot contains variation of both nucleophile and *p*-substituent effect on the benzyl ring. This is the opposite to the Pross and Shaik model in which the transition-state looseness increases as the activation energy increases. The conflicting results imply that either their definition of transition-state looseness is incorrect or that the link between the deformation energy and activation energy of a system is not a simple linear correlation.

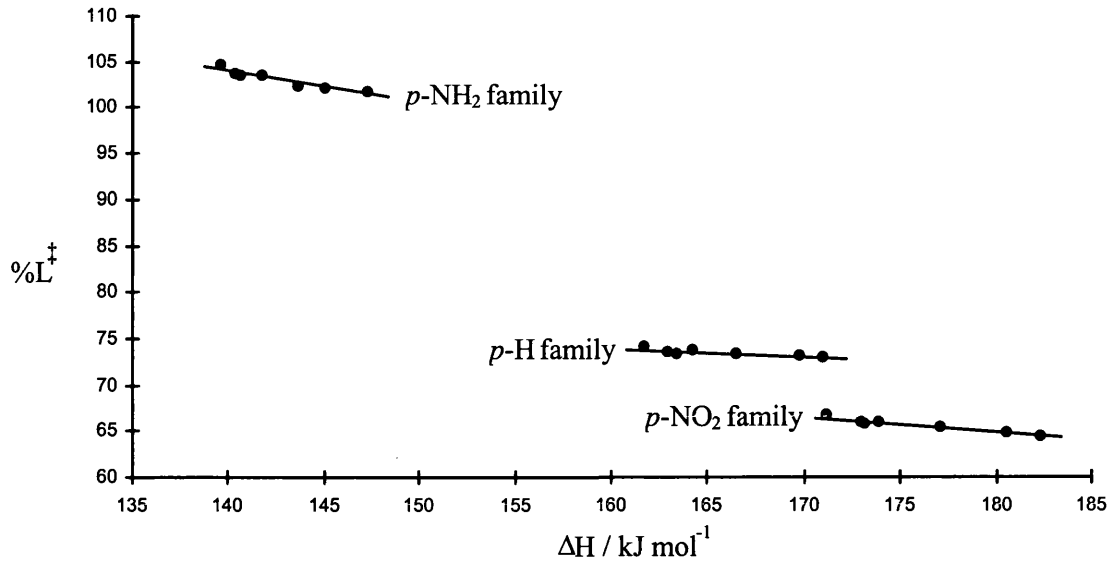


Figure 8.9: A plot of transition-state looseness, %L‡, vs. activation energy, $\Delta H^{\ddagger}$/kJ mol^{-1}.

8.5 Conclusion

The predicted effects of changing substitutional parameters of an S_N2 reaction mechanism by the More O'Ferrall-Jencks and the valence bond configuration mixing model of Pross and Shaik and their comparison with the calculated results are summarised in Table 8.4. In conclusion, the Pross and Shaik model does not correctly predict all of the effects of changing a wide ranging substitutional parameters, whereas the MOFJ model does. The Pross and Shaik model fails dramatically with its prediction of transition-state loosening when the benzyl group is destabilised, whereas the MOFJ model correctly predicts that transition-state tightening would occur. The correlation of barrier height (deformation energy) to transition-state 'looseness' is dependant on their definition of 'looseness'. The Pross and Shaik percentage definition of looseness has come under recent criticism in relating the decreasing value of secondary kinetic isotopes effects to transition-state looseness.[27]

The participation of solvent stabilisation is unclear in both models. To what extent does the solvent stabilises the reactants, products or transition-state? Is there preferential stabilisation of one species over the other? The overall effect of solvation can be mimicked by a continuum method such as COSMO, however for the real structural role of solvent stabilisation to become apparent, a QM/MM investigation would have to be carried out. The Pross and Shaik model is primarily based on the methyl transfer reactions with identical nucleophile and leaving groups. The findings from these

calculations and their relationship with experimental results have been used to explain other, non-related systems. The effect of varying parameters such as leaving group, nucleophile or substituent on a benzyl ring may or may not have the same effect as varying the corresponding parameters in a simple methyl transfer reaction.

Chapter Nine

Bimolecular Transition-states for Simple Alkyl Chlorides

Chapter 9- Bimolecular Transition-states for Simple Alkyl Chlorides

9.1 Introduction

Nucleophilic substitution is one of the most important synthetic reactions in organic chemistry. Hughes and Ingold recognised that a variety of nucleophilic substitution reactions were related by common mechanisms.[32] A logical extension of the calculation of semi-empirical continuum activation energies for the solvolysis reaction for a series of protonated alcohols (see Chapter 4) was the inclusion of explicit solvent molecules. It has been previously shown that this can be achieved by the use of a mixture of molecular mechanics for the solvent and various levels of quantum mechanics for the reaction system.[99] In this study we shall look at the bimolecular solvolysis of a series of simple chloroalkanes in aqueous solution. This will be achieved by using a variety of computational methods whose trends will be compared to available experimental kinetic isotope effects.

9.2 Methods

The calculations are performed using continuum and quantum mechanical/molecular mechanical models. The AM1/COSMO quantum mechanical method is used to describe a continuum model.[11,12] The quantum mechanical/molecular mechanical (QM/MM) model is based on the default AM1 model in CHARMM.[15] The QM/MM models and the transition-state refinement are controlled by an external program that has been developed in our laboratory by A. J. Turner, GRACE.[60] The transition-state from AM1 QM/MM is used to initiate QM/MM modelling with the MP2/6-31G* method. The MP2/6-31G* method is chosen as the starting point because work by Turner has shown that, after an extensive search at various levels of restricted Hartree Fock theory, no saddle point for the aqueous solvolysis of chloromethane could be located. The TIP3P[100] molecular mechanics model is used for all waters except the water involved in the nucleophilic attack, which is treated quantum mechanically. The CAMVIB and CAMISO programs are used to compute the kinetic isotope effects.[28]

9.2.1 GRACE

GRACE is based on a mathematical program whose functions have been oriented to the needs of QM/MM modelling.[101] The main interface of GRACE is via a tool control language which is a structured programming language. It provides a sophisticated 'nerve centre' which uses external codes as well as its own internal algorithms. Visualisation codes are not part of GRACE; however, it can interface with programs such as XMol[102] or RasMol[103] via a series of pdb files which can be read and animated. This makes the visualisation of eigenvectors during saddle point location possible ensuring that the correct mode is followed. The interface with GAMESS-UK[104] operates as part of the *ab initio* QM/MM modelling method in GRACE. When the energy and gradient of a QM/MM system is required, GRACE constructs an appropriate input deck for GAMESS-UK, runs the code and then parses out the energy and gradients from the log file. The molecular mechanics components of the QM/MM function are then constructed via an interface with CHARMM. (See Appendix B for more details on GRACE)

9.2.2 AM1/COSMO

A reaction profile of the S_N2 nucleophilic substitution reaction is computed using the AM1/COSMO method and the approximate saddle point geometry is used initiate a saddle search. Once a transition-state is located, a set of 'valence' internal co-ordinates is then generated using GRACE and CAMVIB which remove the rotational and translational contamination from the hessian before the kinetic isotope effects are computed by CAMISO.

9.2.3 AM1 QM/MM

One approach to model a reaction in solution is to use a combined quantum mechanical (QM) and molecular mechanical (MM) method, Figure 9.1. This 'water droplet' method is split into two regions. One region, the reacting core, contains all the reacting atoms of interest. The atoms in the reacting core are subjected to the quantum mechanical simulation. The surrounding region contains all the atoms outside the reacting core. In this case they are solvent molecules. A convenient way to describe the behaviour of a system being studied with a hybrid potential, as in CHARMM, is to use a notation involving an effective Hamiltonian, H_{eff}.[105] In this formulation, the energy of the

system at a given configuration of the atoms, E, is obtained by solving the time-independent Schrödinger equation, which has the form:

$$H_{eff}\Psi(r_i, R_\alpha, R_M) = E(R_\alpha, R_M)\Psi(r_i R_\alpha R_M) \qquad \text{Equation 9.1}$$

where r_i and R_α stand for the co-ordinates of the electrons and the nuclei in the QM region and R_M for the co-ordinates of the atoms in the MM region, respectively. Ψ is the wavefunction whose square gives the electron density in the system.

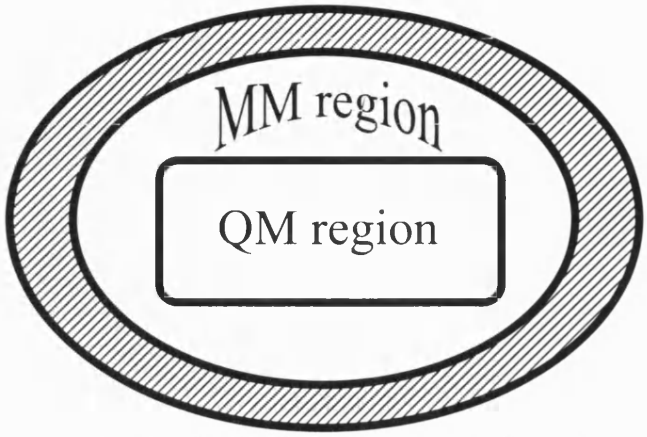


Figure 9.1: A schematic of a simple QM/MM system into its various regions.

The total energy of the system can be written as:

$$H_{eff} = H_{QM} + H_{MM} + H_{QM/MM} \qquad \text{Equation 9.2}$$

where H_{QM} is the quantum mechanical and H_{MM} is the molecular mechanical energy. $H_{QM/MM}$ is the interaction energy between the QM and MM parts of the system which is described by the Hamiltonian, $H_{QM/MM}$. The MM Hamiltonian, H_{MM}, has the interactions between MM atoms only and can be equated to the standard molecular mechanics energy, E_{MM}. This will itself consist of the sum of independent energies, such as the 'bonded' bond, angle and dihedral terms and the non-bonded electrostatic and Lennard-Jones interactions, Equation 9.3.

$$H_{MM} = E_{MM} = E_{bond} + E_{angle} + E_{dihedral} + E_{Coulomb} + E_{Lennard\text{-}Jones} \qquad \text{Equation 9.3}$$

The MM atoms are represented by partial charges and van der Waals spheres centred at the atoms centre, through which the atoms interact at long range. They are linked by harmonic bond and other internal co-ordinates through which they interact at short range and which determine the connectivity of the molecule.

The QM Hamiltonian, H_{QM}, incorporates the nuclear-electron attractions as well as the electron-electron repulsions in the QM system:

$$H_{QM} = -\frac{1}{2}\sum_i \nabla_i^2 + \sum_{ij}\frac{1}{r_{ij}} - \sum_{i\alpha}\frac{Z_\alpha}{r_{i\alpha}} + \sum_{\alpha\beta}\frac{Z_\alpha Z_\beta}{r_{\alpha\beta}} \qquad \text{Equation 9.4}$$

where Z_α is the charge on nucleus a and r_{xy} is the distance between particles x and y. The subscripts (i, j) and (a, b) refer to electrons and nuclei, respectively. CHARMM can either use the AM1[8] or MNDO semi-empirical methods to describe the QM region. The method uses a valence-electron, minimum basis set approximation so that the QM system is comprised of fixed cores made up of the nuclei and inner shell of electrons, plus the valence shell electrons which are treated quantum mechanically. The valence electrons occupy MOs made up of linear combinations of Slater-type orbitals and only the minimum basis set is used. In the cases studied in this thesis, all the reacting atoms in the system are treated QM and MM is used to describe the solvent. This means the term, $H_{QM/MM}$, is due entirely to the non-bonded interactions between the QM and MM atoms. This interaction energy can be written as:

$$\hat{H}_{qm/mm} = -\sum_i\sum_M \frac{q_M}{r_{i,M}} + \sum_\alpha\sum_M \frac{Z_\alpha q_M}{R_{\alpha,M}} + \sum_\alpha\sum_M \left(\frac{A_{\alpha,M}}{R_{\alpha,M}^{12}} - \frac{C_{\alpha,M}}{R_{\alpha,M}^{6}}\right) \qquad \text{Equation 9.5}$$

Where Z represents the charge of the QM atom and R is the distance between electrons i and j. The subscript i in Equation 9.5 refers to a QM electron and the subscript α to a QM nucleus. The subscript M indicates a MM nucleus and q_M is its partial atomic charge. The term describes the QM and MM electron, nuclear and van der Waals interactions. The second and third term in the equation do not involve electronic co-ordinates and so can be calculated in a straight forward manner (i.e. they are constant for a given nuclear configuration).

The TIP3P MM model was used for all solvent water molecules. The TIP3P uses a total of three sites for the electrostatic interaction sites and a rigid water geometry. There are partial positive charges on the hydrogen atoms which are balanced by a negative charge on the oxygen atom. The van der Waals interaction between water molecules is computed using a Lennard-Jones function with a single interaction point per molecule centred on the oxygen atom; no van der Waals interactions involving the hydrogen atoms are calculated.

9.2.3.1 Two zone optimiser

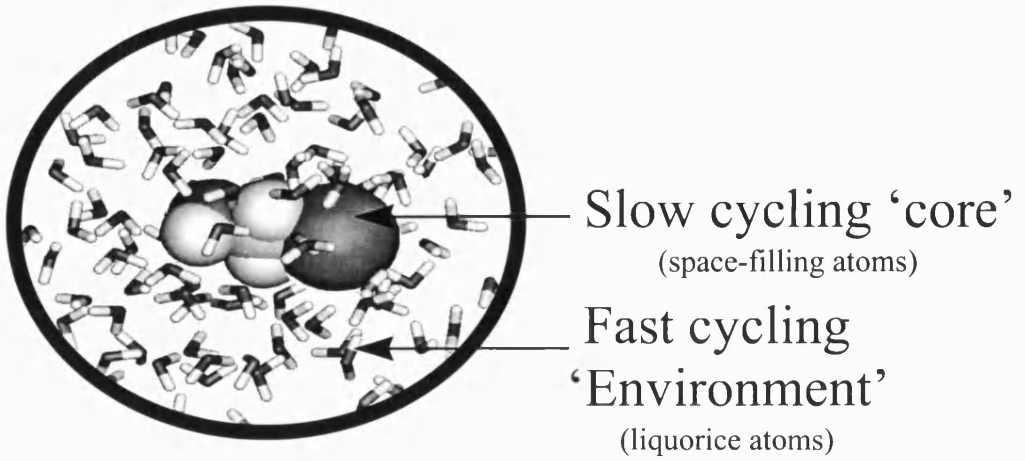


Figure 9.2: A diagram of the two zone optimiser: the fast cycling 'environment' attempts to maintain a gradient of zero while it searches for a stationary point as defined by the slow cycling 'core'.

The optimisations at AM1 QM/MM are divided into two zones, Figure 9.2. One of the zones can be thought of as a fast cycling 'environment' to the second, slow cycling zone which is referred to as the 'core'. In this thesis the 'core' refers to the reaction system i.e. chloroalkane + nucleophile. The algorithm attempts to maintain the gradient of the environment at zero and the potential energy at a minimum. It then searches for a saddle point in the degrees of geometric freedom as defined by the core. Before the energy and gradient used for the core is evaluated, the environment is optimised to a minimum. This means that the number of cycles of the optimiser acting on the environment is much greater than that for the core. Due to the fact that the environment is being constantly minimised, only the hessian matrix for the core needs to be stored and maintained. The ABNR[15] optimiser is used to store a diagonal hessian and a small number of the previous steps and gradient vectors. The ABNR optimiser is chosen due to its residency in CHARMM. Once the transition-state is resolved, its hessian matrix is then computed containing the core and any water molecules which are within 2.5Å of the core. A smaller hessian with just the core atoms is used to start an intrinsic reaction co-ordinate (IRC) calculation in which the environment is continually relaxed.[21] The core atoms follow the mass weighted reaction co-ordinate down from the saddle point to reactants. A few steps of the ABNR optimiser are required to lower the gradient, and the resultant structure is then used to compute a reactant hessian matrix. For completion, a forward IRC to the products is carried out.

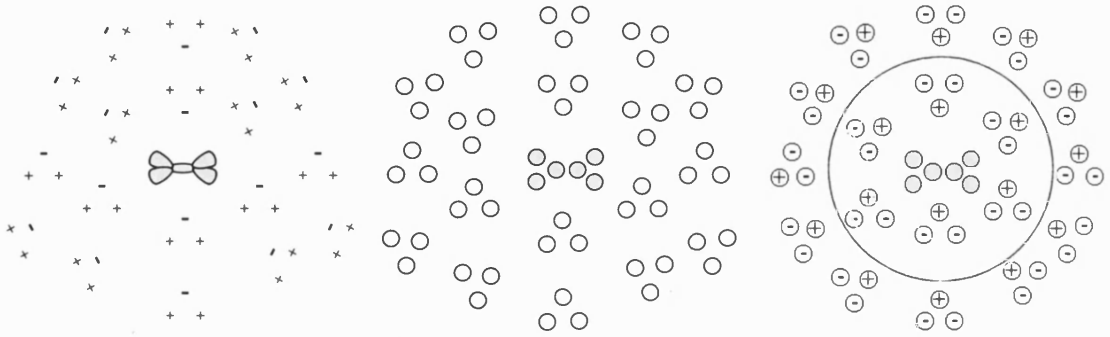
To initially set up the system, the solute is hydrated within a 15Å radius TIP3P water sphere centred on the carbon atoms of the solute molecule. It is then divided into the QM and MM regions: the solute molecules are included into the QM zone, while all the water molecules were treated by MM. The resulting molecular system is a pseudo-sphere of a total amount of ca. 500 water molecules. To obtain an initial guess of the S_N2 transition-state, a reaction profile of the S_N2 nucleophilic substitution is computed and the geometry of the approximate saddle point is then used to initiate a saddle search using a two zone optimiser in GRACE.

The exception to this was for *t*-butyl chloride. After a C–Cl bond separation of approximately 2.9Å is achieved, the chloride would abstract a proton from the cationic *t*-butyl fragment to give $H^+ + Cl^-$ and *iso*-butylene. This problem of proton abstraction from a *t*-butyl carbocation has also been seen by Hartsough and Merz, when investigating the unimolecular solvolysis of *t*-butyl chloride by a QM/MM approach.[106] The elimination reaction has been taken as evidence for elimination proceeding from an ion-pair. [107] The elimination reaction seen by Hartsough and Merz's and our simulation may be an artefact of the limited size of the QM region in the hybrid calculations. Hartsough and Merz's method of stopping the proton abstraction was to constrain the plane of the three methyl carbons to be perpendicular to the C–Cl bond axis throughout the simulation. Our approach to QM/MM modelling is to allow as much geometric freedom as possible for the reacting system and so the constraining of major bonds is considered to be unacceptable in this case. To initiate the saddle search for *t*-butyl chloride, the transition-state geometry derived from AM1/COSMO method is used.

9.2.4 *Ab initio* QM/MM

The located structures for the saddle point from the AM1 QM/MM calculations are used to initiate the *ab initio* QM/MM calculations. It should be noted that due to the computational expense of having 500 TIP3P water molecules, only 200 are used in the *ab initio* calculations. The *ab initio* QM/MM model is controlled by GRACE using the GAMESS-UK package and CHARMM in conjunction with a three zone QM/MM optimiser. The QM/MM potential energy function is generated from three main components; the electrostatic/quantum, bonded/van der Waals and the 'non-polarising'

term, Figure 9.4. The electrostatic/quantum term is made up of contributions for the polarising MM interactions with itself and the quantum atoms, plus the QM atoms interaction with themselves. The Hartree Fock molecular orbital theory with and without second order Møller Plesset perturbation theory is used in this simulation. The bonded/van der Waals component of the QM/MM energy function contains all the bonded MM interactions between the polarising MM atoms with themselves and the van der Waals interaction between the polarising MM atoms with themselves and with the QM atoms.



Electrostatic Van der Waals Non-polarising

Figure 9.3: A diagrammatic model of the QM/MM potentials main components. The electrostatic, van der Waals and non-polarising term.

The van der Waals terms are computed by the MM package. All normal interactions in the MM part of the system, and interactions between the MM section and the QM section are included. The non-polarising contribution to the potential energy function contains all point charge electrostatic interactions and all van der Waals interactions between the non-polarising atoms (outside the line) and the polarising MM and QM atoms (inside the line). It also contains the entire MM force field components for interactions between the non-polarising atoms and themselves. Each QM/MM function evaluation requires a call to GAMESS-UK to acquire the QM terms, a call to GRACE's coulombic algorithm term to remove unwanted coulombic interaction energy and two calls to CHARMM to get the rest of the MM terms.

9.2.4.1 Three zone optimiser

The *ab initio* QM/MM system can be broken down into three zones: a QM, a MM polarising environment and a MM non-polarising environment, Figure 9.4. All the core atoms are treated quantum mechanically. Any water molecules which are within 3.5Å of the core are treated as being in a polarising environment. The atoms in the polarising environment are allowed to 'polarise' the atoms in the QM zone. The remaining atoms are treated as being in an unpolarising environment. The saddle searching was performed using a three zone search algorithm. The semi-empirical QM/MM model located saddle points by only including the QM atoms in the core. If this methodology was continued then the 'core' would contain the atoms for both the QM and the polarising zones. This would significantly increase the size of the hessian matrix, and cause problems for the optimiser in following and maintaining the correct mode.

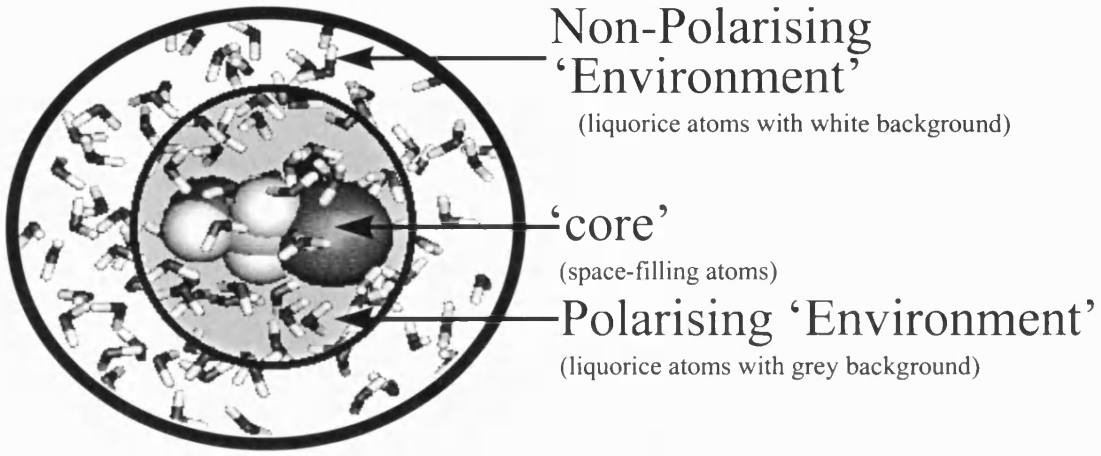


Figure 9.4: A representation of the three zone *ab initio* QM/MM method containing a QM, a MM polarising environment and a MM non-polarising environment.

For the *ab initio* QM/MM optimisation, a third zone is required. This third zone allows the core to contain just the QM atoms and none of the MM polarised environment atoms. The conjugate gradient optimiser by Powell is used to optimise the polarised environment atoms to a minimum for each step of the slow cycling P-RFO optimiser.[108] For each step of the polarised environment optimiser, the non-polarised environment atoms are optimised to a minimum using the ABNR optimiser. Once the saddle point has been located, the intrinsic reaction co-ordinate is computed Hessian matrices are calculated in the same manner as for the AM1 QM/MM calculations however using the three zone instead of two zone QM/MM model.

9.3 Results

The reaction studied by the different solvent models was the bimolecular nucleophilic solvolysis reaction of a series of increasingly substituted chloroalkanes, Equation 9.6, where R = methyl, ethyl, *iso*-propyl and *t*-butyl.

$$RCl + H_2O \rightarrow ROH_2^+ + Cl^- \qquad \text{Equation 9.6}$$

The activation energies, transition-state bond lengths and Pauling bond orders for the different solvent models are given in Tables 9.1-3, together with the α-deuterium, primary carbon and β-deuterium kinetic isotope effects. The α-deuterium kinetic isotope effects reported are per deuterium atom and the β-deuterium kinetic isotope effects are per deuterated methyl group, both at 298.15 K. Due to the computational expense of locating the transition-state for *iso*-propyl and *t*-butyl chloride at the *ab initio* level, transition-state resolution was not fully completed.

Table 9.1: Calculated bonded lengths, bond orders for the transition-state, activation energies and kinetic isotope effect for the studied series at AM1/COSMO.

AM1 / COSMO	Bond Length		Bond Order		$\Delta H^{\ddagger}$ /			
R group	R-Cl / Å	R-O / Å	$n_{R\text{-}Cl}$	$n_{R\text{-}O}$	kJ mol^{-1}	k_H/k_D	k_C^{12}/k_C^{14}	k_{CH_3}/k_{CD_3}
Me	2.295	1.832	0.44	0.53	134.6	0.96	1.104	*
Et	2.419	1.903	0.36	0.47	132.4	0.99	1.100	1.13
^{i}Pr	2.654	2.010	0.24	0.39	125.8	1.08	1.078	1.22
^{t}Bu	3.333	2.035	0.08	0.38	104.9	*	1.046	1.28

Table 9.2: Calculated bonded lengths, bond orders for the transition-state, activation energies and kinetic isotope effect for the studied series at AM1 QM/MM.

AM1 QM/MM	Bond Length		Bond Order		$\Delta H^{\ddagger}$ /			
R group	R-Cl / Å	R-O / Å	$n_{R\text{-}Cl}$	$n_{R\text{-}O}$	kJ mol^{-1}	k_H/k_D	k_C^{12}/k_C^{14}	k_{CH_3}/k_{CD_3}
Me	2.343	1.692	0.40	0.67	179.9	0.94	1.117	*
Et	2.464	1.764	0.33	0.59	149.0	0.98	1.115	1.17
^{i}Pr	2.702	1.842	0.22	0.52	135.8	1.08	1.094	1.25
^{t}Bu	3.108	1.960	0.11	0.43	108.8	*	1.054	1.33

Table 9.3: Calculated bonded lengths, bond orders for the transition-state, activation energies and kinetic isotope effect for the studied series at *ab initio* QM/MM.

Ab initio QM/MM R group	Bond Length R-Cl / Å	R-O / Å	Bond Order $n_{R\text{-}Cl}$	$n_{R\text{-}O}$	$\Delta E^{\ddagger}$ / kJ mol^{-1}	k_H/k_D	k_C^{12}/k_C^{14}	k_{CH_3}/k_{CD_3}
Me (HF/3-21G)	2.022	2.384	0.69	0.21	12.3	0.972	1.095	*
Me (HF/6-31G*)	2.189	2.260	0.52	0.26	51.8	0.984	1.089	*
Et (HF/6-31G*)	2.599	2.115	0.26	0.33	180.7	1.052	1.089	0.98
Me (MP2/6-31G*)	2.182	2.086	0.53	0.35	103.8	0.965	1.107	*
Et (MP2/6-31G*)	2.487	2.093	0.32	0.34	81.2	1.014	1.092	1.06
^{i}Pr (MP2/6-31G*)	2.509	2.055	0.31	0.36	-	-	-	-

$\Delta E^{\ddagger}$: The energy difference between the reactants and transition-state. For *ab initio* QM/MM this represents the difference in relaxation energy of the MM force field and the difference in total electronic + nuclear repulsion energy for the quantum mechanical force field and the QM/MM interaction energy. MP2: 6-31G* basis set used restricted Hartree Fock with MP2 electron correlation on all orbitals.

9.4 Discussion

9.4.1 Experimental results

The experimental α-deuterium kinetic isotope effects for the solvolysis reactions for haloalkanes were measured by Robertson *et al.*[109] The α-deuterium kinetic isotope effects per deuterium atom for a series of bromoalkanes were determined at 298.15 K as 0.96, 0.99 and 1.08 for methyl, ethyl and *iso*-propyl respectively. The α-deuterium kinetic isotope effect for methyl chloride was measured at 0.97 per deuterium atom. The β-deuterium kinetic isotope effect for ethyl bromide (1.03), *iso*-propyl bromide (1.16) and *t*-butyl chloride (1.35) were measured per deuterated methyl group. It is empirically possible to convert kinetic isotope effects for the bromo to the chloro substituted atom. This alters the bromoalkanes experimental values to 1.00 and 1.09 for ethyl and *iso*-propyl respectively, for the pseudo-chloride α-deuterium kinetic isotope effect. The enthalpy of activation for the aqueous solvolysis of methyl chloride had been experimental determined by Robertson as 100.42 kJ mol^{-1}.

9.4.2 AM1/COSMO

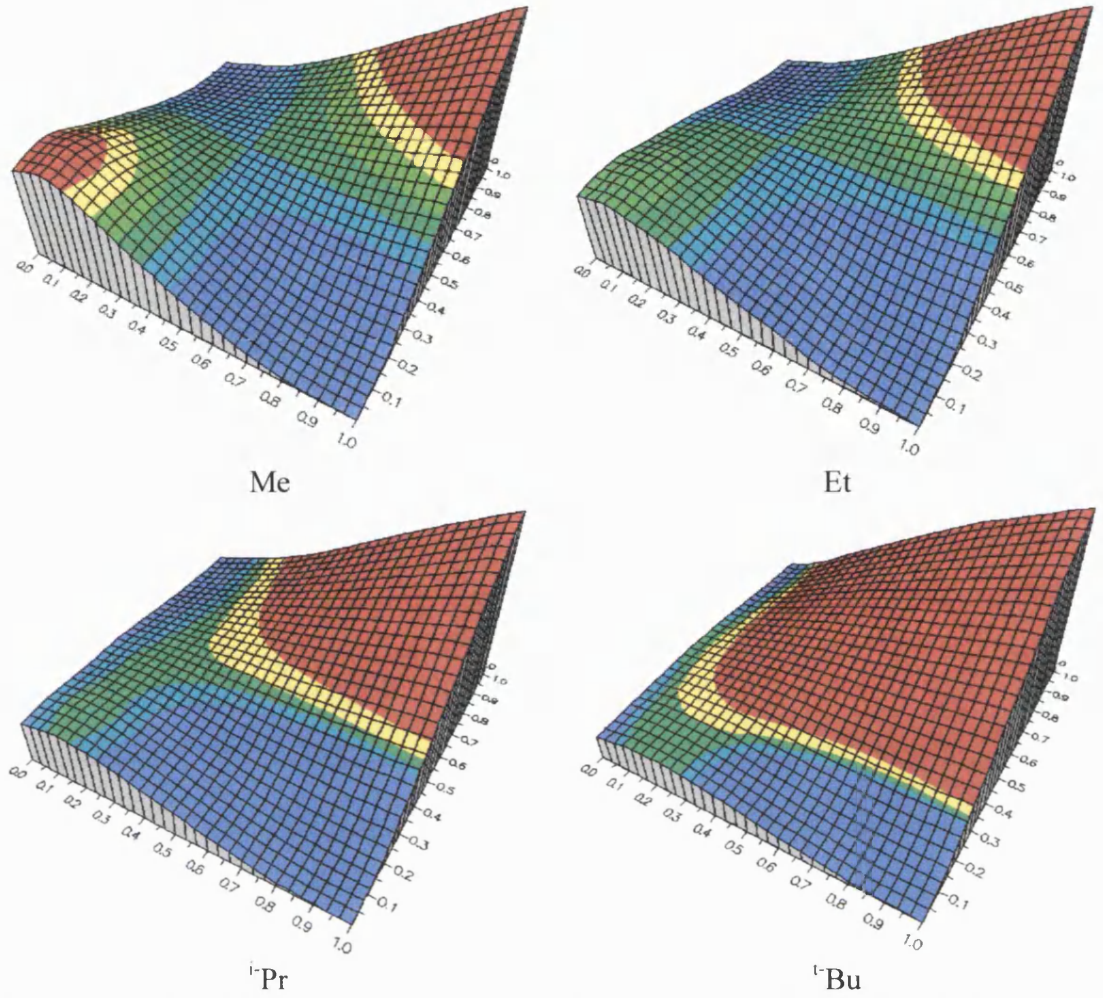


Figure 9.5: AM1/COSMO *in aquo* energy surfaces for the reaction, $RCl + H_2O \rightarrow ROH_2^+ + Cl^-$ where R = Met, Et, ^{i}Pr or ^{t}Bu. Low energies are coloured blue while high energies are coloured red.

The AM1/COSMO energy surfaces shown in Figure 9.5 were produced as described in Chapter 4. The only difference is that the whole surface had to be calculated as the energy surfaces are asymmetrical. The reactants, $RCl + H_2O$, are positioned at the front of the energy surface with the products, $ROH_2^+ + Cl^-$, positioned at the back corner. The left hand side of the energy surface is the dissociative corner, $H_2O + R^+ + Cl^-$, and at the right hand side is the associative corner, $[H_2O–R–Cl]^+$. Due to the asymmetry of the reaction, the surfaces are 'skewed'. From the energy surfaces, it can be seen that the energy of the dissociative corner decreases as the carbocation is stabilised by the addition of methyl groups to the α-carbon. This has the effect of shifting the position of the S_N2 transition-state towards this corner. The S_N1 mechanistic pathway can also be seen on

the energy surfaces. The S_N2 activation energies show a decrease in height as the series is descended with each additional methyl substituent placed on the α-carbon, Table 9.1. In going from methyl to *t*-butyl, the S_N2 transition-state becomes looser.

The α-deuterium kinetic isotope effect for solvolysis reactions are useful criteria of the degree of nucleophilic participation by the solvent in the rate determining step, since typical S_N2 reactions show effects near unity while reactions more S_N1-like in character show effects greater than 1.15.[110] The calculated α-deuterium isotope effects are inverse for methyl and ethyl, and greater than unity for *iso*-propyl. The value for *iso*-propyl approaches a value that has been attributed to an 'exploded' S_N2 transition-state or for a reaction that is classed as having a 'borderline' mechanism. The increase in stability of the carbocation along the series methyl, ethyl, *iso*-propyl and *t*-butyl is associated with an increase in both the inductive effect and hyperconjugation. This is reflected by the increase in the β-deuterium isotope effect as the series is descended. Since a bond to carbon is always broken in S_N1 and S_N2 mechanisms, a carbon isotope effect should be observed.[111] Qualitatively a maximum in carbon isotope effect is observed for a transition-state in which the nucleophile and leaving group are both bonded to carbon with equal strength. The magnitude of the isotope effect at the maximum will also depend on the stiffness of the stretching and bonding motion of the C-H bonds. This is particular important in a 'loose' transition-state where there is a possibility of the inductive effect, conjugation or hyperconjugation. In the AM1/COSMO series, the maximum carbon isotope effect is for methyl which has the tightest transition-state and it then falls through the series as the transition-state becomes looser.

9.4.3 AM1 QM/MM

The activation energy for the S_N2 mechanism decreases as the series is descended with each additional methyl substituent placed on the α-carbon, Table 9.2. In going from methyl to *t*-butyl, the transition-state becomes looser with the C–Cl bond for *t*-butyl becoming greatly elongated 'S_N1-like' in length.[105] However the animation of the transition vectors shows the involvement of both the nucleophile and leaving group, i.e. a bimolecular reaction, Figure 9.6. It should be noted that the involvement of the leaving group does actually decrease as the series is descended and the transition-state becomes more S_N1-like. The C–O bond also increases in length down the series, but to a lesser extent.

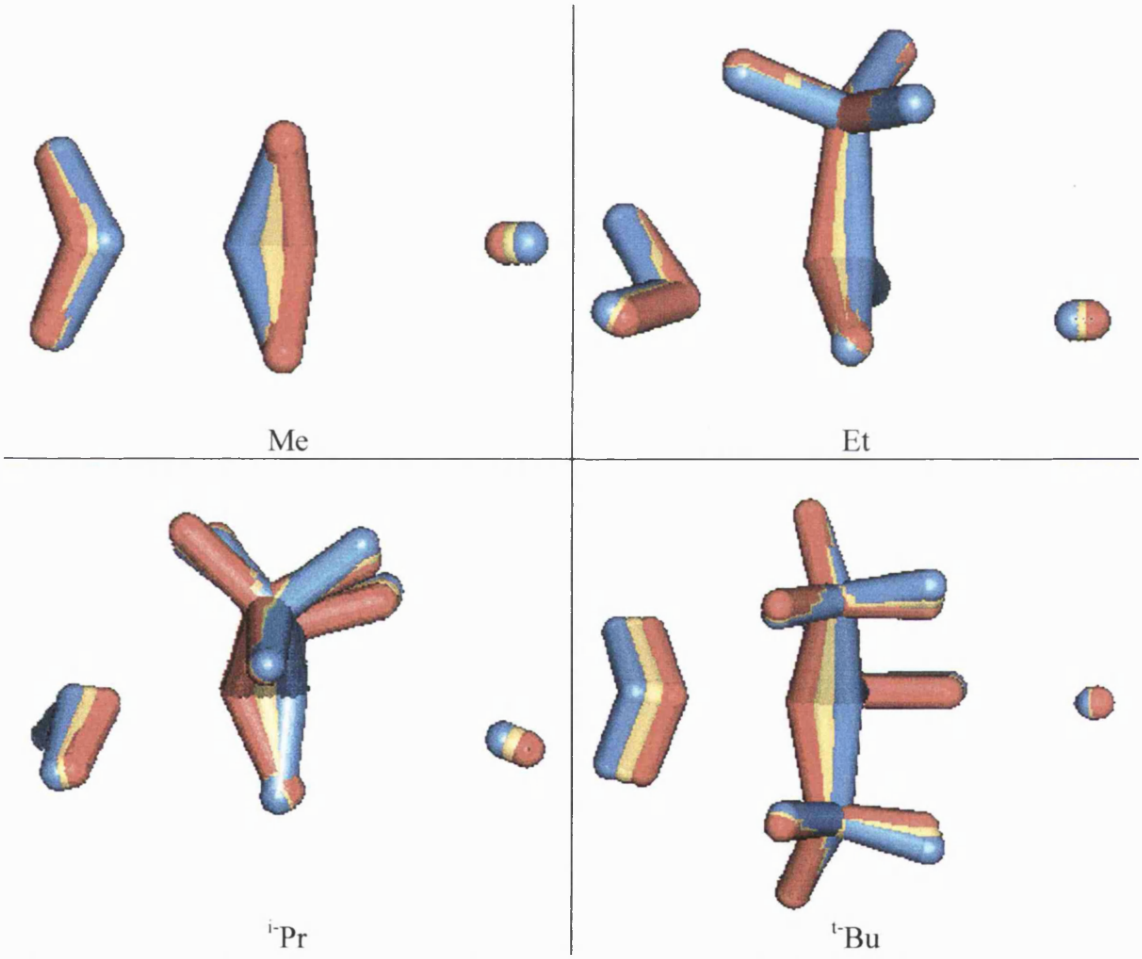


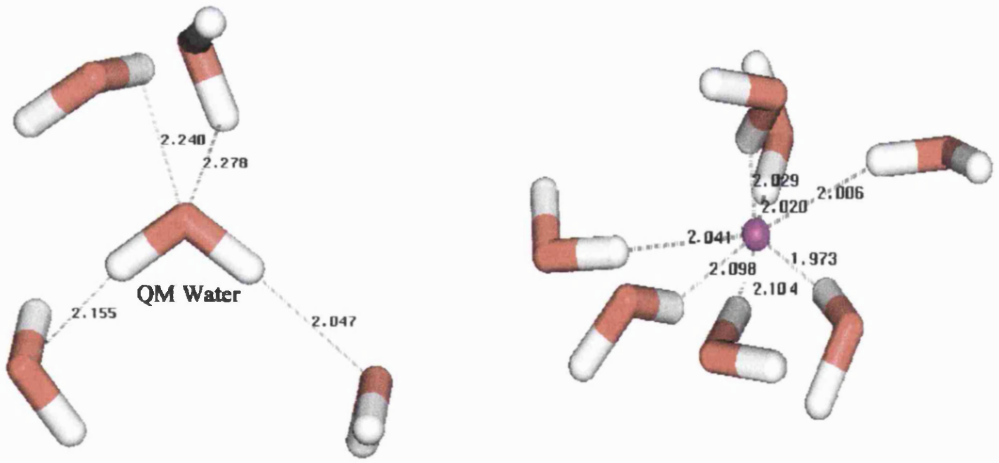
Figure 9.6: The 'animation' snapshots transition vectors at AM1 QM/MM.

The AM1/COSMO energy surfaces shows that the S_N1 as well as the S_N2 pathways coexisted. In the AM1 QM/MM series, the bimolecular transition-state was not specifically targeted for resolution. The AM1 QM/MM transition-states are similar to the

AM1/COSMO structures and would therefore be positioned at a similar 'S_N2' location on the energy surface. The calculated α-deuterium isotope effects for the series are inverse for methyl and ethyl, but greater than unity for *iso*-propyl. The value for *iso*-propyl approaches a value that has been attributed to an 'exploded' S_N2 transition-state or for a reaction that is classed as having a borderline mechanism. The increase in the β-deuterium isotope effect from ethyl to *t*-butyl is a consequence of this increase in hyperconjugation as methyl groups are added. In the AM1 QM/MM series, the maximum carbon isotope effect is seen for methyl which has the tightest transition-state, it then falls through the series as the transition-state becomes looser.

9.4.3.1 Nucleophile and leaving group solvent structures

One of the main advantages of QM/MM over continuum modelling is the fact that it involves specific solvent molecule interactions which can give insight into the structural stabilisation that solvation offers as the reaction proceeds. The figures shown are only snapshots of stationary points. They are not unique stationary points but can each be regarded as one member of a large family or group of similar stationary points. For clarity, the majority of solvent molecules have been removed.



(a) (b)

Figure 9.7: (a) Solvated water nucleophile and (b) the solvated chloride leaving group at AM1 QM/MM.

The structure of the solvated water nucleophile at AM1 QM/MM shows four strong hydrogen bonds to the solvent, Figure 9.7(a). Each of the hydrogens on the QM water creates one strong hydrogen bond while the two lone pairs on the oxygen of the QM

water creates one strong hydrogen bond to each of the solvent molecules. The structure of the solvated chloride leaving group at AM1 QM/MM possesses seven strong hydrogen bonds to the solvent, Figure 9.7(b), as determined by the Cl^-----HOH distances.

9.4.3.2 Reactant complex solvent structure

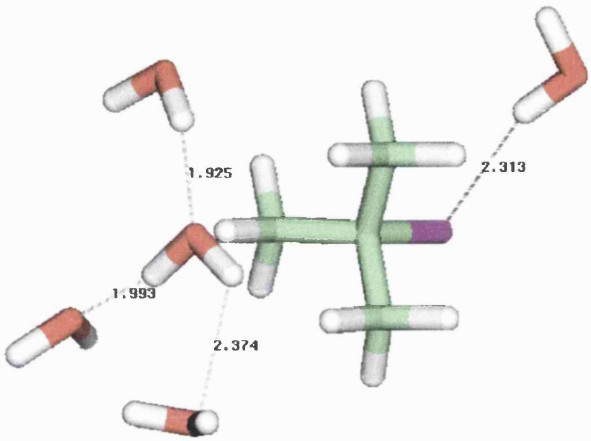


Figure 9.8: The solvated reactant complex of *t*-butyl chloride at AM1 QM/MM.

The reactant complex at AM1 QM/MM has a total of four strong hydrogen bonds associated with it, Figure 9.8. As compared with the solvated water nucleophile, the complexed nucleophile shows that it has lost one water molecule which was donating a hydrogen bond to one of its lone pairs. The alkyl halide has only one 'close' water molecule associated with it, which is a water molecule hydrogen bonded to the chlorine atom. This picture is consistently repeated throughout the series for the reactant complex.

9.4.3.3 Transition-state solvent structures

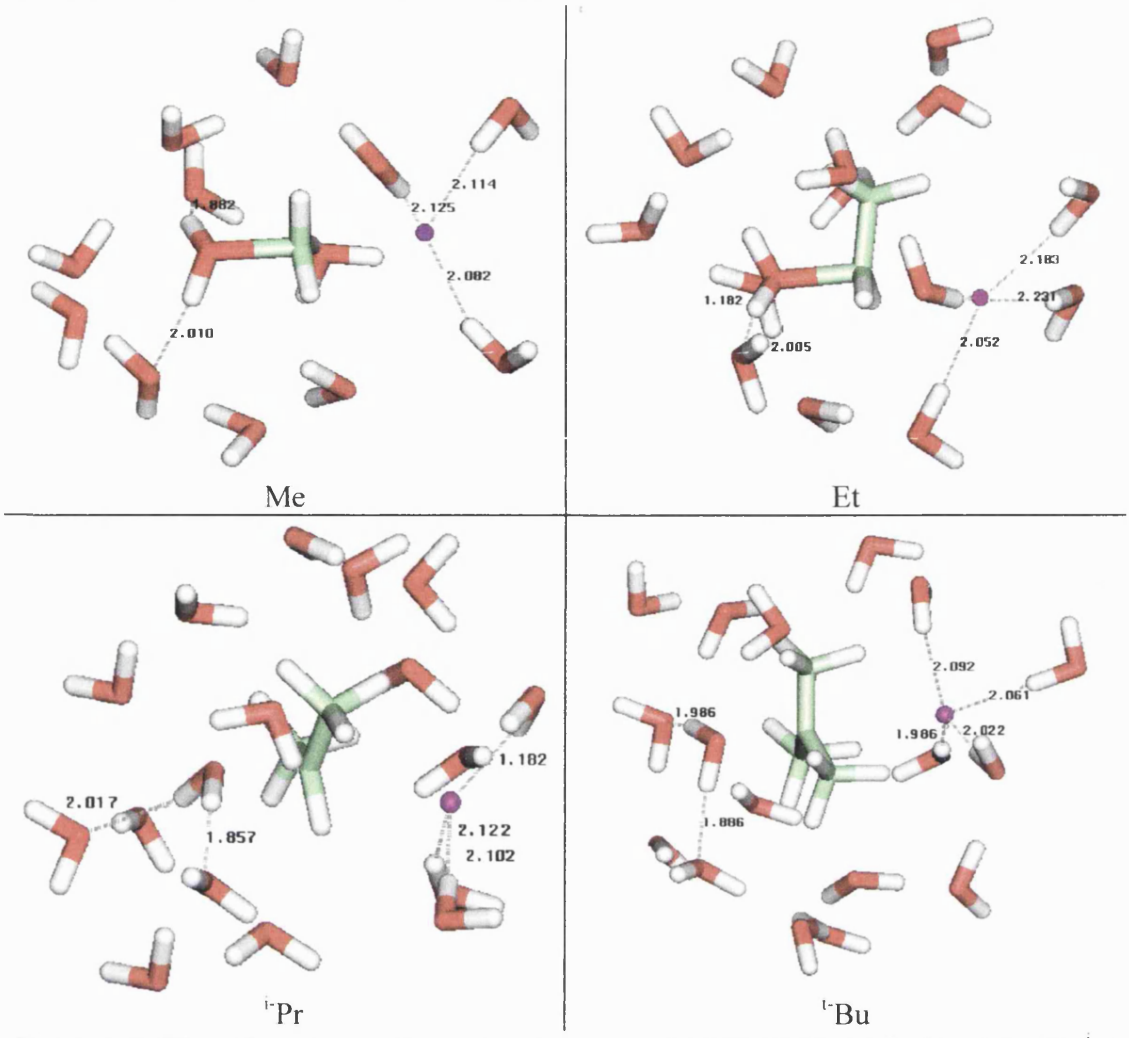


Figure 9.9: The solvated transition-state structures at AM1 QM/MM for R = Me, Et, ^{i}Pr and ^{t}Bu.

The snapshots of the transition-states show all solvent molecules any atom of which lies within 2.8Å of the core (reacting system). The examination of the transition-state structures at AM1 QM/MM show a common orientation of solvent around the nucleophile and leaving group, Figure 9.9. The negatively charged leaving group accepts hydrogen-bonds to three (or four in the case of *t*-butyl) water molecules, whereas the positively nucleophile remains a hydrogen-bond donor to two waters but is not involved as a hydrogen-acceptor. What is apparent from the structures is that there are no close waters associated with the methyl substituents on the α-carbon atom. There appears to be a linkage between the nucleophile and leaving group by a system of 'water chains' in the first solvation shell, Figure 9.10.

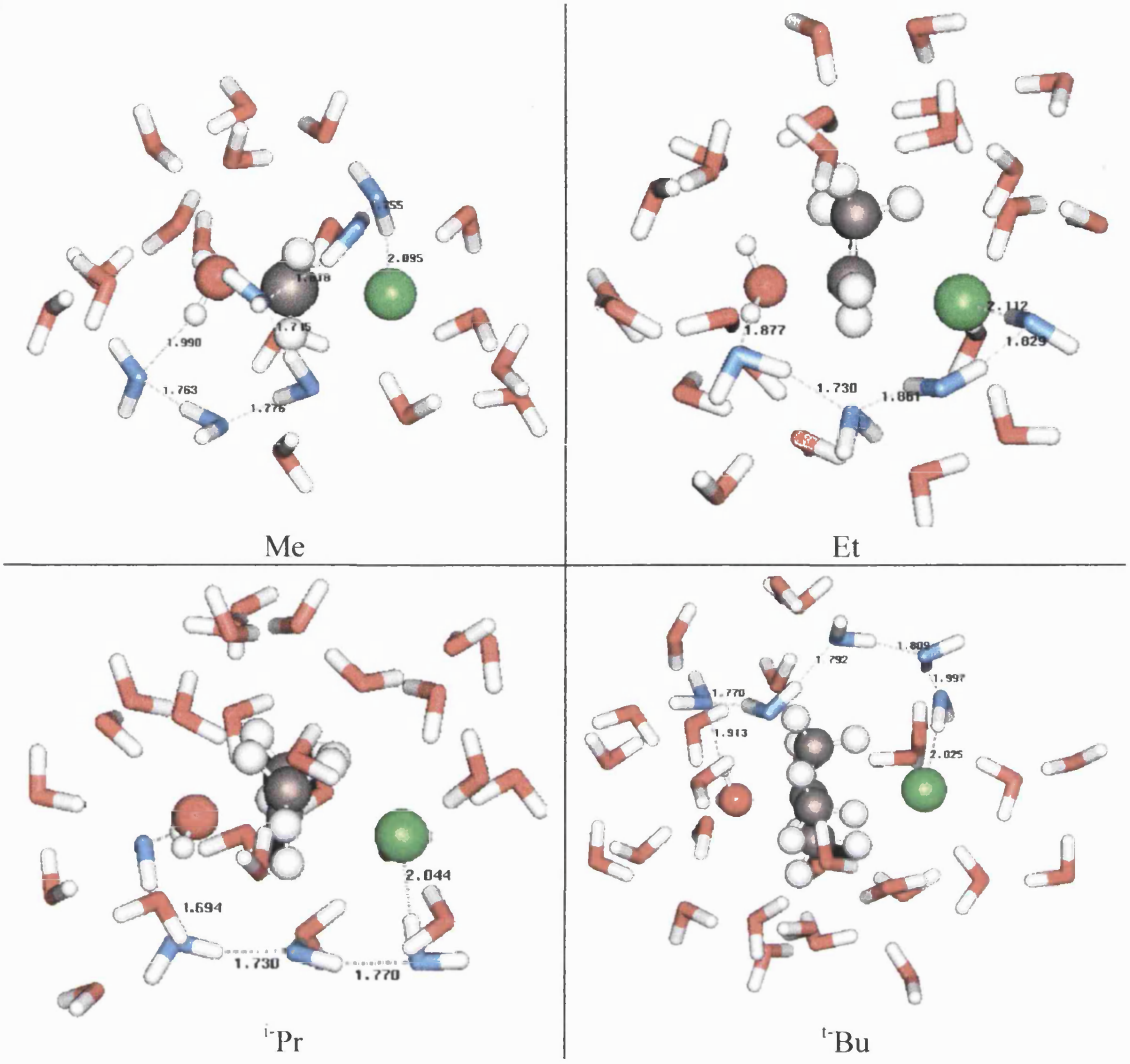


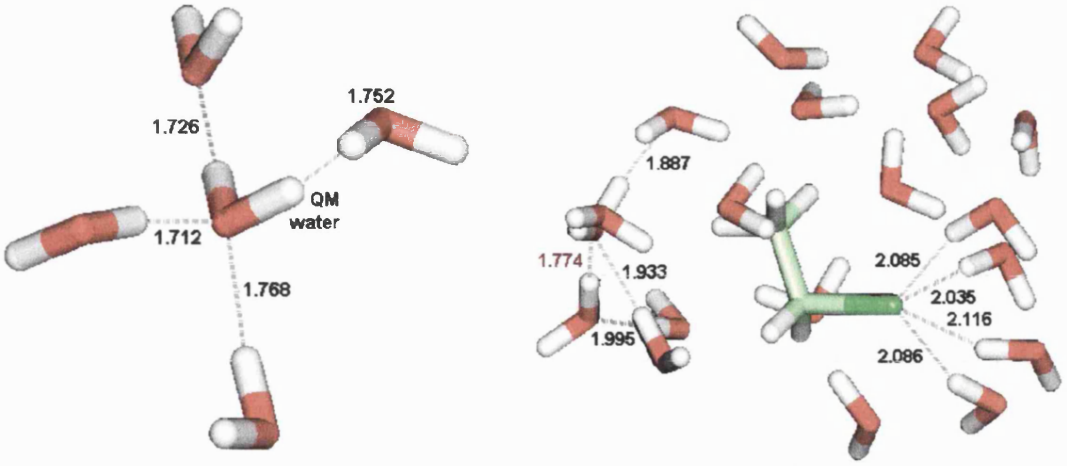
Figure 9.10: The solvated transition-state structures at AM1 QM/MM for R = Me, Et, ^{i}Pr and ^{t}Bu depicting water chains (blue coloured solvent waters).

9.4.4 *Ab initio* QM/MM

Turner[60] had stated that the MP2/6-31G* method was chosen as the starting level of theory because after he carried out an extensive search at various levels of restricted Hartree Fock theory, no saddle point could be located. However by 'backtracking' using the MP2/6-31G* structure as the initial starting point, it was possible to locate the transition-states at the 3-21G and 6-31G* Hartree Fock level. The activation energy for the S_N2 mechanism increases as each additional methyl substituent is placed on the α-carbon, Table 9.3, at the HF/6-31G* level whereas, common with the semi-empirical methods, it decreases for the series at the MP2/6-31G* level. The comparison of the experimental activation energy for the aqueous solvolysis of methyl chloride is in good agreement with the MP2 calculated result whereas the Hartree Fock barriers are to low.

In going from methyl to *iso*-propyl, the transition-states become looser. The extent to which the C–Cl bond has dissociated at the transition-state is less in the *ab initio* QM/MM than either of the semi-empirical methods. Whereas the extent to which the C–O bond has dissociated at the transition-state is more in the *ab initio* QM/MM than either of the semi-empirical methods. The calculated α-deuterium isotope effect is inverse for methyl and greater than unity for ethyl whereas the semi-empirical methods both gave ethyl an inverse α-deuterium isotope effect. The MP2/6-31G* calculated α-deuterium kinetic isotope effects give the same qualitative trend for the series as experimental results. The MP2/6-31G* calculated kinetic isotope effect for bimolecular solvolysis of methyl chloride is in perfect agreement with experiment and the calculated result for the rest of the series is also in very good agreement with the pseudo isotope effects. The HF/6-31G* calculated isotope results show the poorest correlation with the experimental results. It should be noted that the HF/3-21G calculated result gives a good agreement with the experimental isotope effect however the calculated barrier height is far to low.

9.4.4.1 Nucleophile and Reactant complex solvent structure



a b

Figure 9.11: (a) Solvated water nucleophile and (b) the solvated reactant complex at MP2/6-31G* QM/MM.

The structure of the solvated water nucleophile at MP2 QM/MM shows four strong hydrogen bonds to the solvent, Figure 9.11a. Each of the hydrogens on the QM water creates one strong hydrogen bond while the two lone pairs on the oxygen of the QM water create one strong hydrogen bond to each of the solvent molecules. As compared with AM1 QM/MM, the nucleophile to solvent hydrogen bonds are much shorter in distance at the MP2 QM/MM level of theory. The reactant complex at MP2 QM/MM possesses four strong hydrogen bonds, Figure 9.11b. As compared with the solvated water nucleophile, the complexed nucleophile again shows that it has lost one water molecule which was donating a hydrogen-bond to one of its lone pair. The alkyl halide possesses a greater number of water molecules associated with its chlorine atom as compared to AM1 QM/MM.

9.4.4.2 Transition-state solvent structures

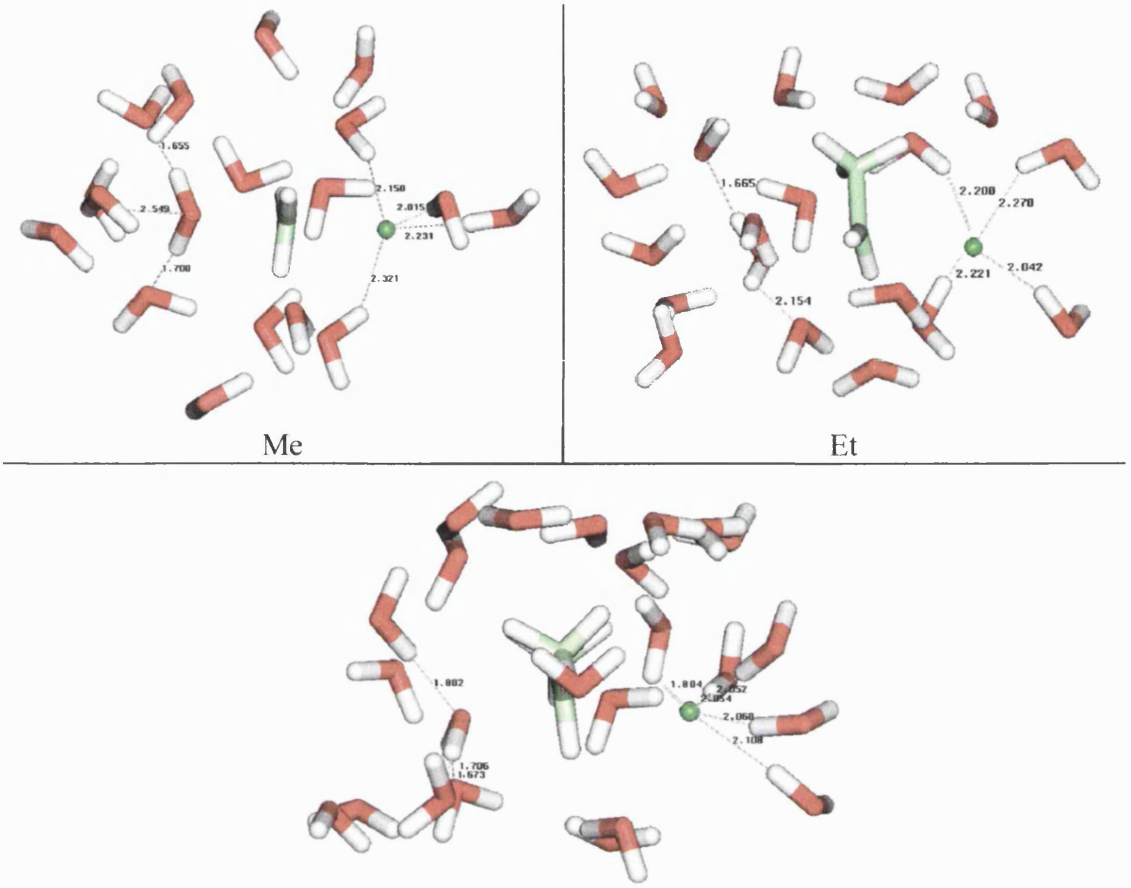


Figure 9.12: The solvated transition-state structures for R = Me, Et and ^{i}Pr at MP2/6-31G* QM/MM.

The examination of the MP2 QM/MM transition-state structures shows that the number of solvent molecules encompassed within a 2.8Å radius of the core has significantly increased, Figure 9.12. For the methyl and ethyl transition-states, the leaving group has four 'close' hydrogen bonding water molecules surrounding it whilst the nucleophile has three. Whereas for *iso*-propyl, the leaving group has five 'close' hydrogen bonding water molecules surrounding the chloride anion.

9.5 Conclusion

The close similarity of the results obtained by use of the COSMO continuum method to those obtained by the QM/MM method, seems to indicate that in certain cases the AM1/COSMO results can be used to predict or even initiate the AM1 QM/MM model. This was demonstrated in the location of the *t*-butyl transition-state which otherwise would have resulted in an elimination and not a substitution reaction. The COSMO and AM1 QM/MM methods are in close agreement with one another in structure and kinetic isotope effect. The reaction modelled at the *ab initio* (MP2/6-31G*) QM/MM level are in excellent agreement with the available experimental kinetic isotope effects and activation energy.

The main advantage of the QM/MM over continuum methods is the insight into the role solvent stabilisation that they present as demonstrated by the changing solvent co-ordination number of the nucleophile as the reaction proceeds. The Hartree Fock level of theory gave poor results as compared to the MP2 and semi-empirical results. Despite the much greater computational cost of the MP2/6-31G* QM/MM method (e.g. 12 months of continuous calculations on a dedicated DEC-alpha workstation for a single S_N2 transition-state!) as compared to semi-empirical, it would appear that for the reactions studied, its use is justifiable in order to produce results in agreement with experiment.

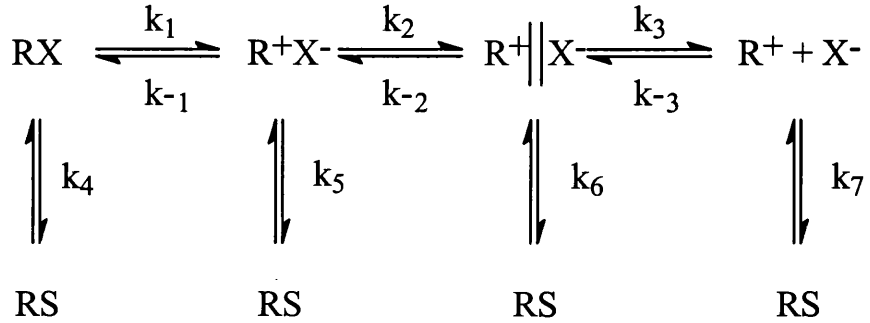


Figure 9.13: The unimolecular pathway depicting the possible roles of ion-pairs in a mechanistic pathway.

The AM1/COSMO energy surfaces reveal the possibility of alternative pathways for nucleophilic substitution. The work carried out in this simulation has only included the S_N2 and not the S_N1 mechanistic pathway. The elucidation of the S_N1 pathway would lend itself to comparison with the results obtained in this S_N2 study. The S_N1 pathway according to Winstein, Figure 9.13, would be best studied by a QM/MM method as it could show any ion-pairing that may occur as the leaving group dissociates.[112]

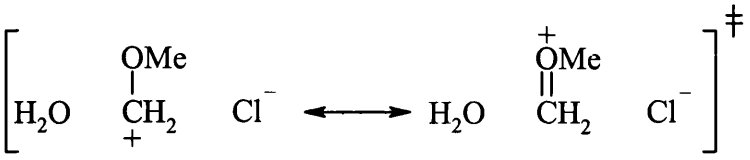
Chapter Ten

The Solvolysis of Methoxymethylchloride

Chapter 10 - The Solvolysis of Methoxymethylchloride

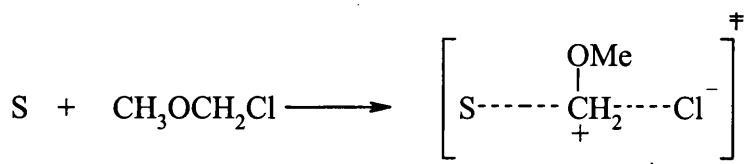
10.1 Introduction

Methoxymethylchloride is known to decompose with rapidity in solvolytic and S_N2 reactions.[113] The kinetic isotope effect for the solvolytic reaction has been experimentally measured and interpreted in various conflicting ways. The methoxymethyl group is an interesting group to study as sterically it resembles a primary halide whereas electronically, it resembles a tertiary halide due to the possibility of transition-state charge stabilisation, Equation 10.1.

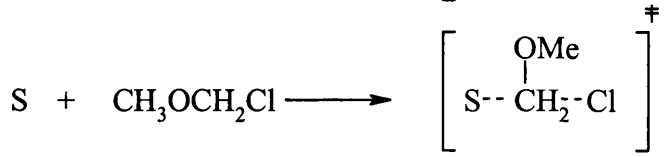


Equation 10.1

Hughes and Ingold postulated, but never proved, that the aqueous solvolysis of the primary halide, methoxymethyl chloride, would be found to belong to the unimolecular category.[34] Experimental evidence for the rate of ethanolysis of methoxymethylchloride in the presence of excess ethoxide ion suggests that the reaction is bimolecular in nature.[114] In a study by Jones and Thornton, they assumed that the difference between an S_N2 and S_N1 mechanism was the 'tightness' of the transition-state. They measured the α-deuterium isotope effect for solvolysis of methoxymethylchloride to be 1.24 ± 0.08 per deuterium atom in 2-propanol at 273.15K. The large magnitude of this isotope effect would usually be associated with a 'loose', S_N1 transition-state, Equation 10.2, whereas if the value was inverse or close to unity, it would be associated with a 'tight', S_N2 transition-state, Equation 10.3. They concluded from their study that the transition-state for methoxymethylchloride is like that for the solvolysis of *t*-butyl chloride. That is, the transition-state is 'loose' with a great deal of carbocation character even though sterically there is a possibility of a tight, S_N2 transition-state.



Equation 10.2



Equation 10.3

S = Solvent

On the investigation of the bimolecular substitution reaction on acetals by Kirby,[115] it was concluded that there was a good deal of evidence that nucleophilic substitutions at centres with strongly electron-donating substituents are characterised by large α-deuterium isotope effects, for S_N2 as well as S_N1 mechanisms. This infers that caution must be taken when attributing kinetic isotope effects to mechanisms. Unlike Jones and Thornton, they attributed the transition-state to be an 'exploded' S_N2 transition-state due to the variance of the kinetic isotope effect when the nucleophile varied. They described the 'exploded' transition-state as either an unusually open transition-state for an S_N2 displacement reaction or as an oxocarbenium ion that is stabilised by weak interactions with the nucleophile and leaving group.[116] The large isotope effect is used to support the idea of an open, exploded transition-state with a large amount of oxocarbenium ion character which has freedom for the bending motion for the C-H bonds of the formaldehyde moiety.

The valence bond configuration mixing model,[97] discussed in Chapter 8, was used by Shaik to investigate the transition-state for methoxymethylchloride reaction.[94] The conclusions drawn from their model was that the transition-state will involve substantial carbocation character due to the stability of the triple-ion configuration, $[N:^-(CH_3OCH_2)^+:X^-]$, induced by the methoxy group. The relationship between molecular distortion and activation energy plays an important role in the Pross and Shaik model. If a reaction is slow, it will require a large amount of molecular distortion, relative to the reactants, whereas if the reaction is fast, it does not. The solvolysis reaction of methoxymethylchloride reaction is reported to be faster than the reaction for methyl case which is known to have a 'tight' transition-state. Therefore according to the model, the transition-state will be tighter than for the corresponding unsubstituted substrate.

10.2 Methods

The aim of this simulation is to investigate the aqueous solvolysis of methoxymethylchloride, Equation 10.4, which will be achieved by using a variety of computational methods. The location of saddle points for the bimolecular solvolysis of methoxymethylchloride in aqueous solution using semi-empirical continuum and semi-empirical QM/MM modelling is undertaken. The calculated isotope effects are

compared to the experimental kinetic isotope effect and the structural information gained is used to indicate whether the transition-state is 'tight' or 'loose'.

$$H_2O + CH_3OCH_2Cl \longrightarrow CH_3OCH_2\overset{+}{O}H_2 + Cl^- \qquad \text{Equation 10.4}$$

The AM1/COSMO[11,12] model is used for the continuum treatment of solvation. The QM/MM model is based on the default AM1 model in CHARMM[15] with the subsequent transition-state refinement controlled by GRACE.[60] The TIP3P molecular mechanics model is used for all waters except the water involved in the nucleophilic attack which is treated quantum mechanically. A detailed description of the computational methods employed can be read in Chapter 9. The CAMVIB and CAMISO programs are used to compute the kinetic isotope effects.[28]

10.3 Results

The calculated results for the solvolysis of methoxymethylchloride from the various computational methods are given in Table 10.1. The transition-state bond lengths (Å) and Pauling bond orders for the carbon-chlorine (C–Cl), carbon-oxygen (C–O) and the α-carbon-methoxy carbon bond lengths, (C-OMe) are given as well as the C-OMe bond length for the reactants; also given are the enthalpy of activation and α-deuterium kinetic isotope effects per deuterium atom at 275.15 K. The Pauling bond orders are calculated with the proportionality constant, c, equal to 0.6 for the σ-bonds and 0.3 for the π-bonds.

Table 10.1: Calculated bonded lengths and bond orders for the S_N2 transition-states for the solvolysis of methoxymethylchloride together with the calculated activation energies, kJ mol^{-1}, and α-deuterium kinetic isotope effect.

Method	TS C–Cl/Å	TS C–O/Å	Reactant C–OMe/Å	TS C–OMe/Å	Bond Orders n_{C-Cl}	n_{C-O}	n_{C-OMe}	$\Delta H^{\ddagger}$	k_H/k_D
AM1/COSMO	2.515	1.934	1.400	1.327	0.30	0.45	1.28	58.4	1.03
AM1 QM/MM	2.579	2.032	1.400	1.308	0.27	0.38	1.36	53.7	1.04
AM1 QM/MM	3.298	3.567	1.400	1.285	0.08	0.03	1.46	66.0	1.25

10.4 Discussion

In their study, Jones and Thornton measured the α-deuterium isotope effect for solvolysis of methoxymethylchloride to be 1.24 ± 0.08 per deuterium atom in 2-propanol at 273.15 K. They also measured the enthalpy of activation, $\Delta H^{\ddagger}$, and entropy of activation, $\Delta S^{\ddagger}$, in 2-propanol, to be 44.8 ± 2.5 kJ mol^{-1} and -120.5 ± 6.7 J K^{-1} mol^{-1},

respectively. The AM1/COSMO method resolved a bimolecular transition-state for the solvolysis of methoxymethylchloride, Table 10.1. The transition-state possesses a great deal of bonding from the nucleophile and leaving group towards the α-carbon which gives rise to a tight transition-state, and an α-deuterium kinetic isotope effect associated with a S_N2 mechanism, a value close to unity. The calculated enthalpy of activation is approximately 10 kJ mol^{-1} too high compared with the experimental value.

The AM1 QM/MM method resolved two sets of bimolecular transition-states for the solvolysis of methoxymethylchloride. From the bond orders, one transition-state can be regarded as being 'tight' while the other is 'loose'. The 'tight' transition-state possesses a great deal of bonding from the nucleophile and leaving group towards the α-carbon which gives rise to an α-deuterium kinetic isotope effect associated with a typical S_N2 mechanism. The 'loose' transition-state possesses little bonding from the nucleophile or leaving group towards the α-carbon which gives rise to an α-deuterium kinetic isotope effect associated with an S_N1-like mechanism. The calculated loose KIE of $k_H/k_D = 1.25$, is in very good agreement with the experimentally determined result of $k_H/k_D = 1.24 \pm 0.08$. However, the activation energy for the tighter transition-state is much closer to experiment than that for the loose one. The double bond character of the α-carbon-methoxy carbon bond, C-OMe, Equation 10.1, for the loose transition-state is reflected in the greater bond order for that bond. The QM/MM model resolved two bimolecular transition states whereas the continuum model only resolved one. This is because the continuum model can not describe specific solvent-solute interactions that stabilise a transition-state and so it only the tighter of the two transition-states was resolved.

10.5 The examination of solvent structures

One of the main advantages of QM/MM over continuum modelling is the fact that it involves specific solvent molecule interactions which can give insight into the structural stabilisation that solvation offers as the reaction proceeds. The snapshots show all solvent molecules any atom of which lies within 2.8Å of the core (reacting system). It should be noted that these are 'snapshots' and not unique transition-state structures.

10.5.1 Reactant complex structure

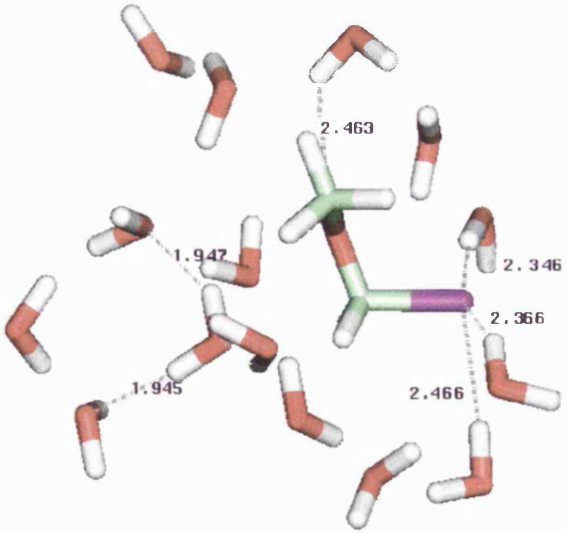


Figure 10.1: The solvated methoxymethylchloride and nucleophile reactant complex at AM1 QM/MM.

The reaction complex at AM1 QM/MM has three hydrogen bonds associated with the nucleophile; three 'loose' hydrogen bonds associated with the chlorine of the reactant and one hydrogen bond associated with the oxygen on the methoxy group, Figure 10.1.

10.5.2 Transition-state structures

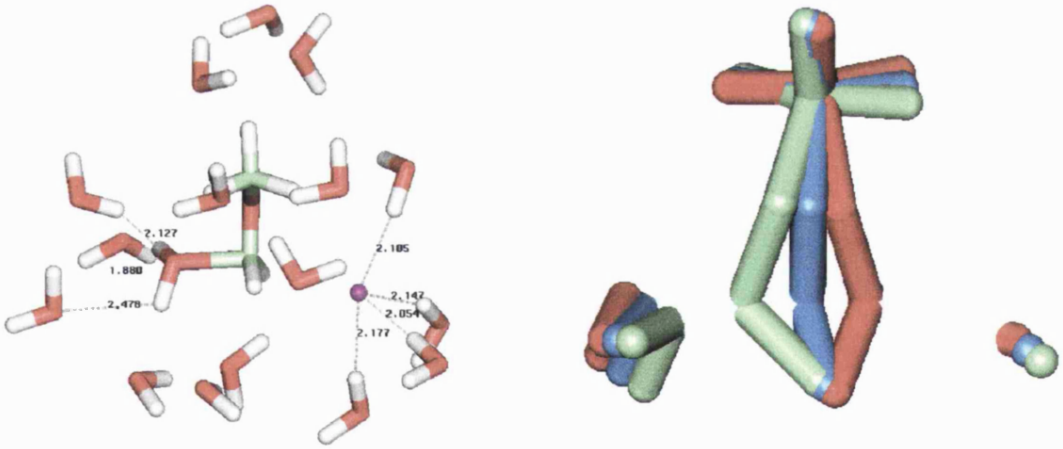


Figure 10.2: The 'tight' transition-state for the solvolysis of methoxymethylchloride at AM1 QM/MM, its 'animated' snapshot.

The 'tight' transition-state at AM1 QM/MM possesses three hydrogen bonds associated with the nucleophile and four with the leaving group which have become 'tighter' as compared to the reactant-complex while the hydrogen bond associated with the oxygen on the methoxy group has disappeared, Figure 10.2. The 'tight' animated transition-state

motion vector shows the involvement of both the nucleophile and leaving group in the transition-state.

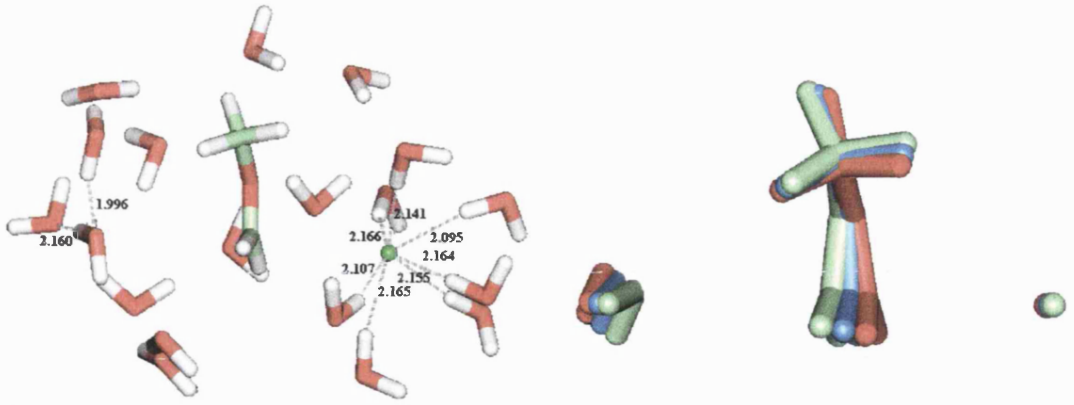
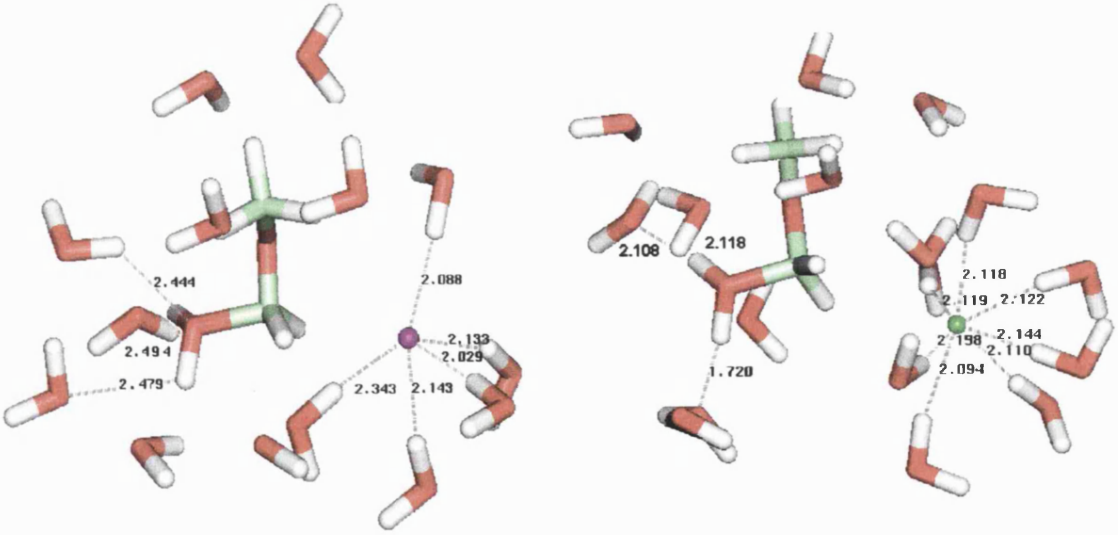


Figure 10.3: The 'loose' bimolecular transition-state structure for the solvolysis of methoxymethylchloride at AM1 QM/MM, its 'animated' snapshot.

The 'loose' transition-state at AM1 QM/MM, possesses three hydrogen bonds associated with the nucleophile and seven hydrogen bonds associated with the leaving group, an increase of three, which have become 'tighter', Figure 10.3. The 'loose' bimolecular animated transition-state motion vector shows that the nucleophile is mainly involved with only slight participation of the leaving group. However, the reaction is still bimolecular.

10.5.3 Product complex structure



(a) tight (b) loose

Figure 10.4: The solvated protonated alcohol and leaving group structure at AM1 QM/MM from the resultant forward IRC paths for (a) the tight and (b) the loose transition-states.

The protonated alcohol that resulted from the IRC for the tight transition-state has three water molecules associated with the product complex water molecule, Figure 10.4a. These are less tightly bound than in the transition-state. The leaving group now has five waters surrounding it, Figure 10.4a, but no water molecules between itself and the product. The protonated alcohol that resulted from the IRC for the loose transition-state has three water molecules associated with the product complex water molecule, Figure 10.4b. Again, these are less tightly bound than in the transition-state. The leaving group now has eight waters surrounding it but no water molecules between itself and the product.

10.6 Conclusion

The AM1/COSMO model resolved a transition-state that was close to the experimental enthalpy of activation but whose calculated KIE was not. The AM1/TIP3P model resolved two transition-state; one tight and the other loose. The loose transition-state was close to the experimental isotope effect. Even though sterically there is a possibility of a tight, S_N2 transition-state, the AM1 QM/MM study shows that the transition-state is 'loose' with a great deal of carbocation character.

The QM/MM model resolved two bimolecular transition states whereas the continuum model only resolved one. As the continuum model can not describe specific solvent-solute interactions that stabilise a transition-state and so it only the tighter of the two transition-states was resolved. The inclusion of specific solvation interaction to describe reaction mechanisms is an important feature of QM/MM modelling. Great care must be taken when initially setting up the system. Achieving a 'realistic' solvent packing in the first shell is of vital importance when setting up the system.

Chapter Eleven

Transition-state Structural Refinement with GRACE and CHARMM: Modelling of Glycoside Hydrolysis in Aqueous Solution Using a Combined Quantum/Classical Method

Chapter 11- Transition-state Structural Refinement with GRACE and CHARMM: Modelling of Glycoside Hydrolysis in Aqueous Solution Using a Combined Quantum/Classical Method

11.1 Introduction

Glycoside hydrolyses are important bio-organic processes whose mechanisms are of considerable interest to physical organic chemists. Glycoside hydrolysis teeters on the brink between a stepwise and concerted mechanism. A hybrid quantum/molecular mechanical simulation of a glycoside hydrolysis model in aqueous solution was under taken by Barnes and Williams.[99] They used the CHARMM QM/MM program to obtain an energy surface showing two distinct reaction pathways.[15] The comparison of experimental data with their calculated isotope effects indicated that the preferred mechanism was stepwise. Barnes and Williams obtained their transition-states by a grid searching methodology which could be thought of as 'unrefined' transition-states.

The multifunctional program GRACE[60] which was developed by our group employs the 'eigenvector-following' method for location of transition structures in association with a subset-Hessian approach (see Chapter 9 for more detail). The use of GRACE allows the location of 'true' or 'refined' saddle points for a proposed mechanisms as characterised by a single imaginary vibrational frequency. These 'refined' saddle points are used in this chapter to calculate kinetic isotope effects for comparison with experimental results to elucidate the mechanism, and to re-evaluate the results of Barnes and Williams.

The basic mechanistic chemistry of the non-enzymic hydrolysis of simple pyranosidic and furanosidic O-glycosides was established by the mid-70's.[117] Evidence exists suggesting that the hydrolysis of N-glycosidic bonds occurs with a considerable oxycarbenium character in the ribose ring at the transition-state. Experimental evidence for the existence of oxycarbenium character was based on a large α-deuterium secondary effect for the acid catalysed solvolysis of inosine and adenosine.[118] The oxycarbenium ion is indicative of a stepwise or 'exploded' concerted transition-state. For the oxycarbenium ion to exist, it must be stabilised internally or by the surrounding

solvent. It is therefore critical that the solvent is described by a discrete method that allows a detailed description of the solvent structure.

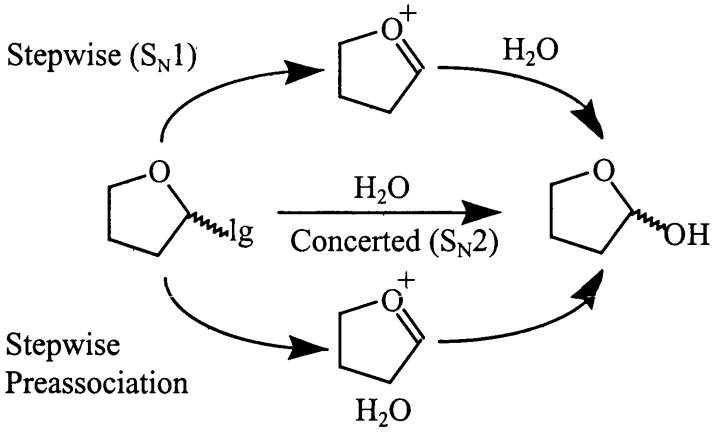


Figure 11.1: Pathways for the hydrolysis of simple glycosides.

A simple glycoside may undergo hydrolysis by three mechanisms, Figure 11.1. The stepwise $D_N + A_N$ (S_N1) mechanism involves rate determining heterolysis of the aglycone leaving group, forming a glycosyl cation which is then subject to nucleophilic attack. The concerted A_ND_N (S_N2) mechanism involves a bond to the nucleophile being formed as the bond to the leaving group is broken. The stepwise $D_N{*}A_N$ (S_N1) pre-association mechanism is similar to the $D_N + A_N$ (S_N1) one with the nucleophile appearing in the rate determining step only as a 'spectator'. Schramm *et al.* measured experimentally kinetic isotope effects for the acid-catalysed hydrolysis of adenosine monophosphate, Figure 11.2, with isotopic substitution at four positions.[119] They concluded from their measured kinetic isotope effects, Table 11.2, that the reaction mechanism was stepwise in nature.

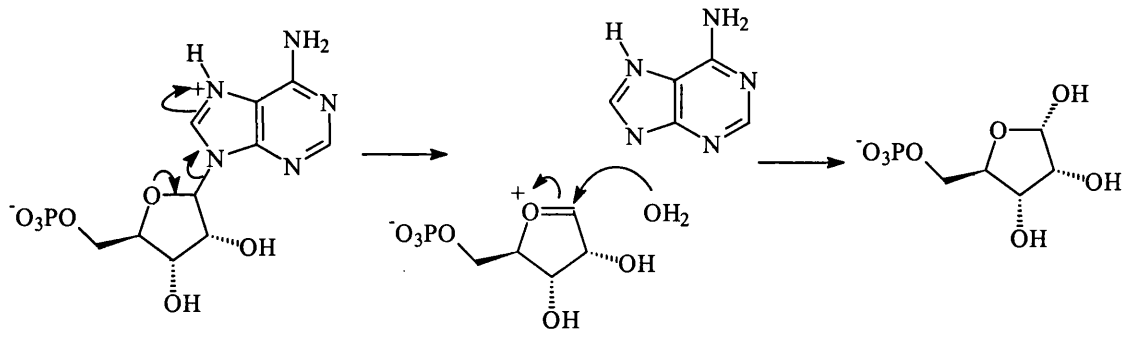
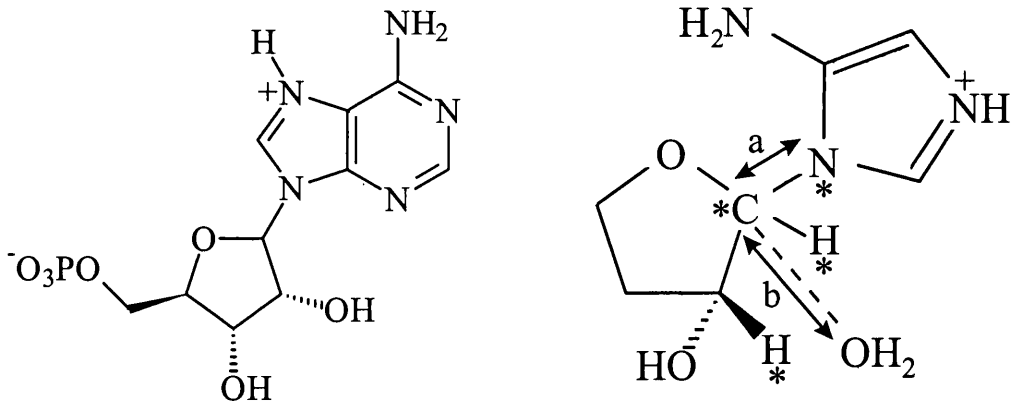


Figure 11.2: The hydrolysis of adenosine monophosphate.



11.3a 11.3b

Figure 11.3: Adenosine monophosphate and its computational model; (a) represents the carbon-nitrogen leaving group distance, (b) represents the carbon-oxygen nucleophile distance and the (*) represents the various points of isotopic substitution.

To reduce the computational expense, a 'model' system of the acid-catalysed hydrolysis of adenosine monophosphate was used, Figure 11.3b. Barnes and Williams used the CHARMM QM/MM program to obtain the energy surface shown in Figure 11.4. The model structure was used to calculate each point on the grid. The points were formed by pairs of values for the making bond (C-Nu) to the nucleophilic water molecule and the breaking bond (C-Lg) to the aglycone leaving group at 0.2Å intervals. The transition-state structures were located accurately by a fine grained (0.02Å) grid search around the approximate saddle points. The topology of the resulting surface revealed a sharp saddle point leading from the reactant valley into a shallow bowl, corresponding to an intermediate, from which the product valley was reached by traversing a second broad saddle region; this suggests a stepwise $D_N{*}A_N$ pre-association mechanism. However the surface also shows the presence of a second, separate pathway leading directly to the products across the ridge from the reactants valley thereby suggested a concerted D_NA_N mechanism. They concluded the possibility existed for both the stepwise $D_N{*}A_N$ and concerted A_ND_N mechanisms. The agreement between calculated and experimental kinetic isotope effects indicated that the preferred mechanism was stepwise.

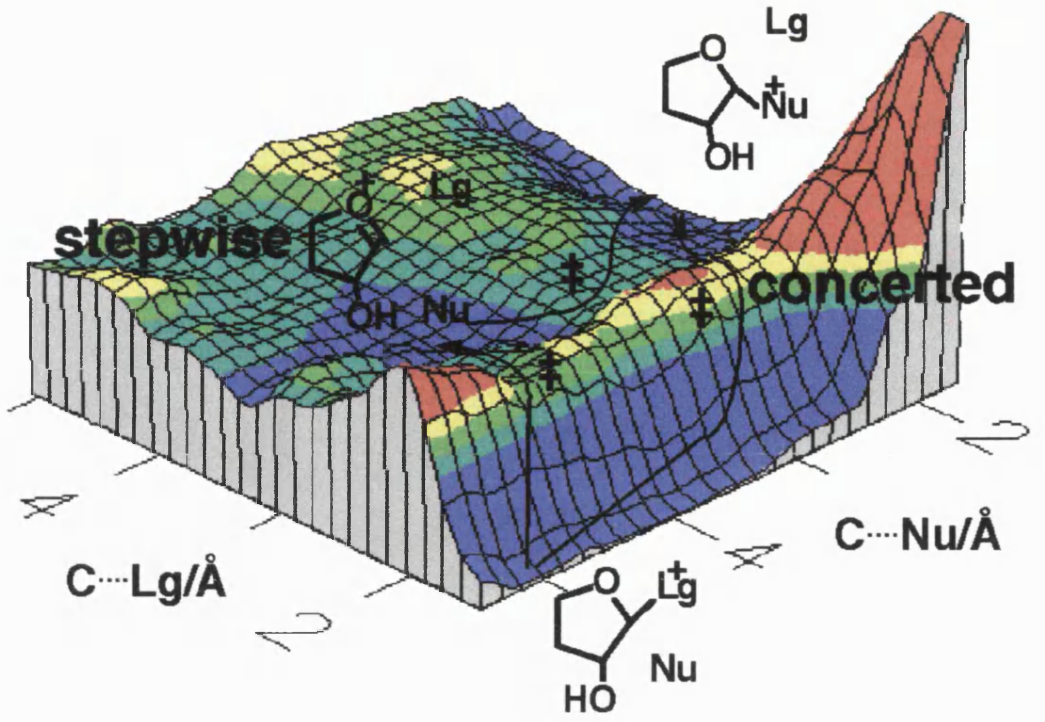


Figure 11.4: The QM/MM energy surface for the model of the acid-catalysed hydrolysis of adenosine monophosphate (Nu = water nucleophile; Lg = aglycone leaving group).

The aim of this simulation is to locate 'true' saddle points for the proposed mechanisms for the model of the acid-catalysed hydrolysis of adenosine monophosphate characterised by a single imaginary vibrational frequency. The calculated isotope effects are compared to experimental values in an attempt to elucidate the mechanism that relates to the experimentally obtained results.

11.2 Methods

A hybrid QM/MM computational study for the model of the glycoside hydrolysis in aqueous solution was undertaken. The study involves the acid-hydrolysis of the model compound embedded in a 15Å radius of ca. 500 water molecules, Figure 11.5. The reacting system is treated by the AM1 semi-empirical molecular orbital method whilst the solvent water molecules are described by the TIP3P empirical potential.[100] Approximate transition-states are obtained using the CHARMM program and then refined by GRACE. The kinetic isotope effects are calculated using the CAMVIB and CAMISO programs.[28]

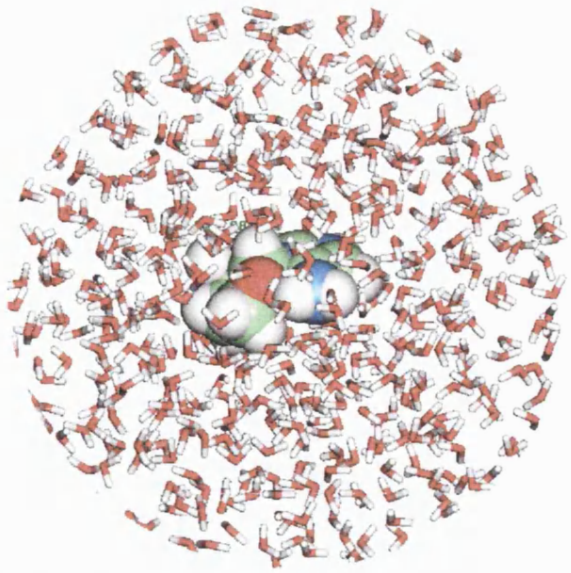


Figure 11.5: A snapshot of the adenosine monophosphate model embedded in a 15Å radius of ca. 500 water molecules.

11.3 Results

Table 11.1: Calculated bonded lengths and Pauling bond orders for the transition-state and the enthalpy of reaction for the acid-catalysed hydrolysis of adenosine monophosphate model.

Mechanism		Bond Length C-N / Å	Bond Length C-O / Å	Bond Order n_{C-N}	Bond Order n_{C-O}	Relative Energy kJ/mol^{-1}
Concerted		2.081	2.363	0.38	0.24	158.2
Stepwise	TS1	2.146	2.827	0.34	0.11	171.6
Stepwise	INT	2.610	2.357	0.16	0.24	143.1
Stepwise	TS2	2.735	2.132	0.13	0.35	167.0

The transition-states for the various mechanisms were obtained and refined by GRACE, Table 11.1. The transition-state bond lengths (Å) and Pauling bond orders for the carbon-nitrogen (C-N) leaving group distance and carbon-oxygen (C-O) nucleophile distance for the various mechanisms are given. The Pauling bond orders are calculated as described in Chapter 4 using the reactant bond length for the equilibrium bond length and a value of 0.6 for the constant c. The relative energies are calculated as the energy difference between the reactant complex and each of the species. For the stepwise mechanism, the energies for the first transition-state (TS1), the intermediate (INT) and

the second transition-state (TS2) are given. The kinetic isotope effects are calculated at 323.15 K, Table 11.2. The experimental and the 'grid' search results for the various isotopic substitutions are also given.

Table 11.2: Experimental kinetic isotope effect for the acid-catalysed hydrolysis of adenosine monophosphate in 2 mol dm^{-3} HCl at 50°C and the QM/MM calculated results for the model.

		QM/MM GRID[99]		QM/MM GRACE		
Isotope Effect	Experimental	Concerted	Stepwise	Concerted	Stepwise TS1	Stepwise TS2
1° ^{14}C	1.044 ± 0.003	1.044	1.045	1.101	1.042	1.039
1° ^{15}N	1.030 ± 0.002	1.028	1.028	1.042	1.024	1.022
2° β-^{2}H	1.077 ± 0.002	1.071	1.074	1.123	1.093	1.068
2° α-^{3}H	1.216 ± 0.004	1.164	1.211	1.321	1.235	1.186
rms. error		0.026	0.003	0.064	0.013	0.016

11.4 Discussion

The GRACE hybrid QM/MM computational study of the model for glycoside hydrolysis in aqueous solution resolved various transition-states for the concerted and stepwise mechanisms. The concerted 'animated' transition-state shows the involvement of both the nucleophile and leaving group in the transition vector, Figure 11.6. The stepwise TS1 'animated' transition-state shows the main involvement of leaving group and little motion of the nucleophile in the transition vector while the stepwise TS2 'animated' transition-state shows the main involvement of nucleophile and little motion of the leaving group.

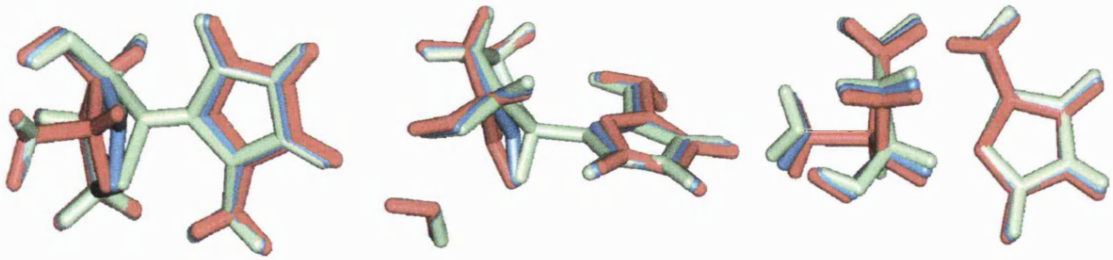


Figure 11.6: The 'animated' transition-state of the glycoside hydrolysis for the concerted and stepwise mechanisms.

As Barnes and Williams' QM/MM energy surface showed, two reaction pathways are possible. From the calculated barrier heights, the concerted mechanism is favoured over the stepwise mechanism by 13.3 kJ mol^{-1}. The concerted transition-state does not resemble an 'exploded' transition-state such that can be seen for the solvolysis of D-glucopyranosyl derivatives or methoxymethyl chloride.[113,120] The bonding to the leaving group in the stepwise mechanism is much looser than that found in the concerted mechanism, whereas the bond to the nucleophile is much shorter. The effect of the water molecule in the stepwise mechanism is initially to stabilise the transition-state (TS1) by forming a strong hydrogen bond to hydroxyl group on the glycosyl ring. The proximity of the water molecule to the formed oxycarbenium intermediate (INT) is then ideal for the subsequent nucleophilic attack to the glycosyl cation (TS2).

A test of any theoretically determined structure is how well it reproduces experimental results. The isotopic substitutions offer information about the change in bonding of that atom in the transition-state. The calculated kinetic isotope effect for each mechanism is given in Table 11.2 for isotopic substitution at four separate positions. The experimental and the calculated CHARMM grid search kinetic isotope effects are also given. The carbon and nitrogen isotope effects are larger for the concerted mechanism due to the fact that the bonding to the leaving group is greater in the concerted transition-state than the stepwise one. Both the secondary α-tritium and the secondary β-deuterium effect for the concerted mechanism is greater than that for the stepwise mechanism. These results are contrary to the conventional expectation that the isotope effect for a concerted mechanism is close to unity. If correct, this would question the validity of assigning mechanisms to reactions solely on their kinetic isotope effects. Schramm *et al.* assigned the stepwise mechanism to this reaction due to the large values of their experimental kinetic isotope effects.

The secondary β-deuterium effect is related to the relative hyperconjugation of the transition-state. The value for the concerted mechanism is greater than that for the stepwise mechanism because of the larger carbon-nitrogen bond dissociation in the concerted transition-state i.e. greater oxycarbenium character. Barnes and Williams' grid

calculated results closely resembled the experimental kinetic isotope effects differing only for the concerted secondary α-tritium effect. This was the criterion by which Barnes and Williams' concluded that the preferred mechanism was stepwise. The GRACE/CHARMM calculated kinetic isotope effects for the stepwise mechanism are much closer to the experimental values than that calculated for the concerted mechanism suggesting that the preferred mechanism is indeed stepwise, despite the relative values of the calculated barriers for the two pathways.

11.5 The examination of solvent structures

One of the main advantages of QM/MM over continuum modelling is the fact that it involves specific solvent molecule interactions which can give insight into the structural stabilisation that solvation offers. The figures shown are only snapshots of stationary points. They are not unique stationary points but can be regarded as members of a large family or group of similar stationary points. A slight reorganisation of a single solvent molecule may perturb the system so that a new transition-state structure results. This 'new' transition-state may or may not be similar in energy and structure to the 'old' transition-state.

11.5.1 Reactant complex structure

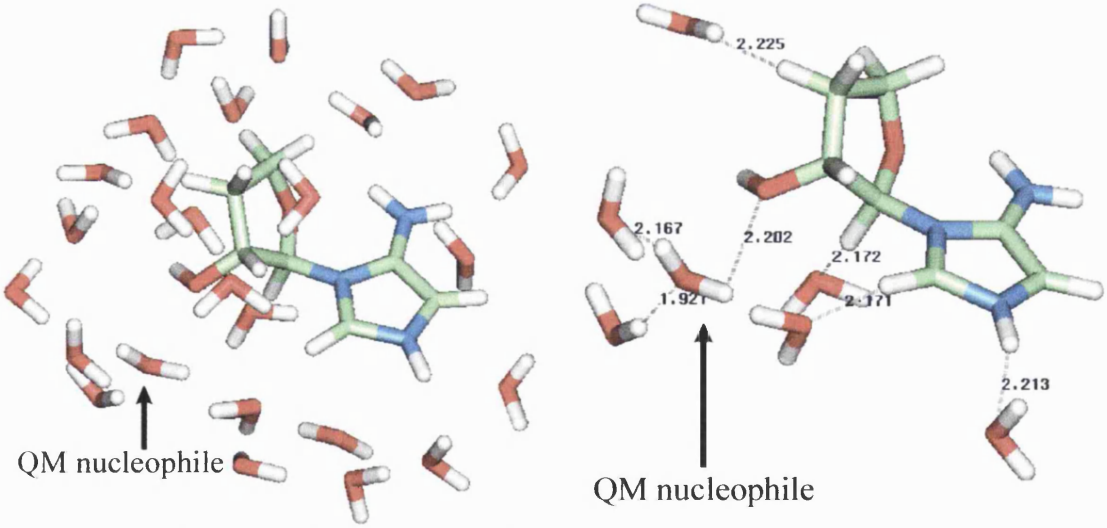


Figure 11.7: The solvated reactant complex at AM1 QM/MM.

The nucleophile in the solvated reactant complex possesses three strong hydrogen bonds; one to the hydrogen, another to its oxygen and one to the hydroxyl oxygen of the glycosyl group, Figure 11.7.

11.5.2 Transition-state structures

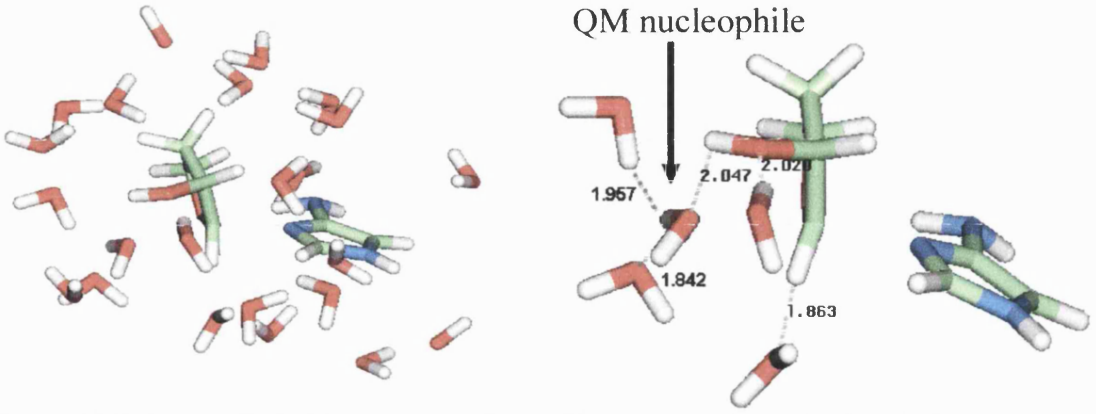


Figure 11.8: A snapshot of the concerted transition-states' first solvation shell.

At the concerted transition-state, each of the hydrogens on the QM nucleophile creates a strong hydrogen bond to the solvent while one of its lone pairs on the oxygen accepts a strong hydrogen bond from the hydroxyl on the glycosyls' ring. The other lone pair is orientated towards the reacting α-carbon centre, Figure 11.8. The hydroxyl group possesses two strong hydrogen bonds; one from the nucleophile and the other from a solvent molecule to the hydroxyls' oxygen.

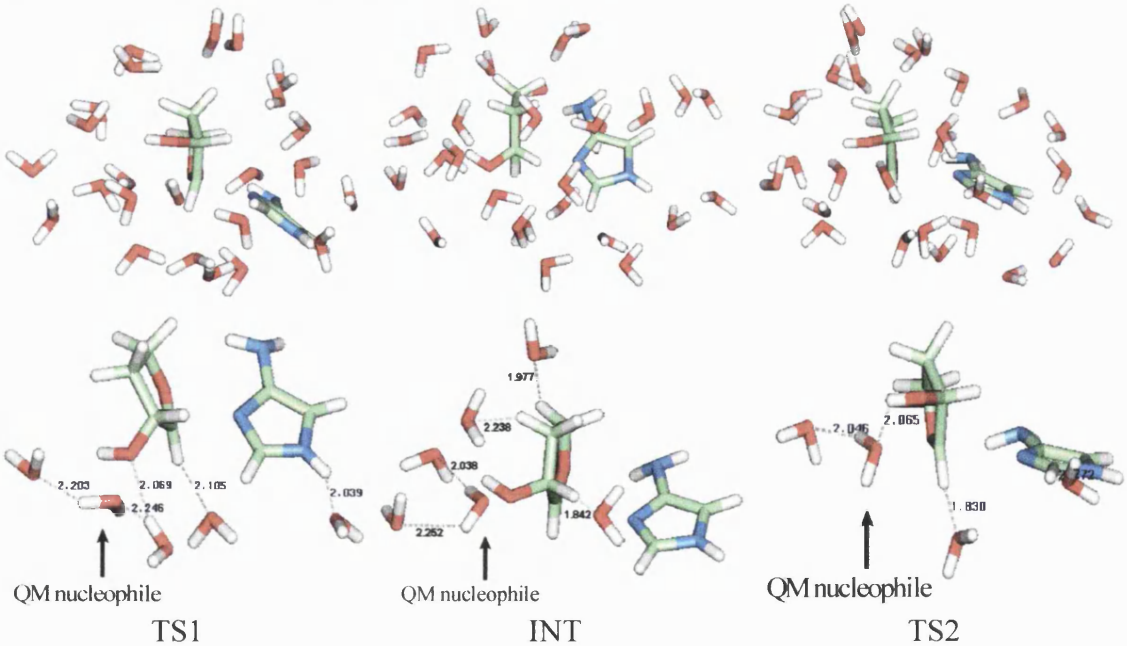


Figure 11.9: A snapshot of the stepwise mechanisms' stationary points.

At the stepwise transition-state (TS1) the attacking nucleophile possesses a strong hydrogen bond towards the hydroxyl group on the glycosyl cation and two others to the surrounding 'bulk' solvent, Figure 11.9. The density of the first solvation shell is greatest at the intermediate stationary point (INT) due to the fact that it requires the greatest solvent stabilisation. This is reflected in the tighter hydrogen bonds from the surrounding solvent molecules for the intermediate stationary point. As the reaction progresses, the solvation of the nucleophile decreases in the act of becoming bonded to the α-carbon.

11.5.3 Product complex structure

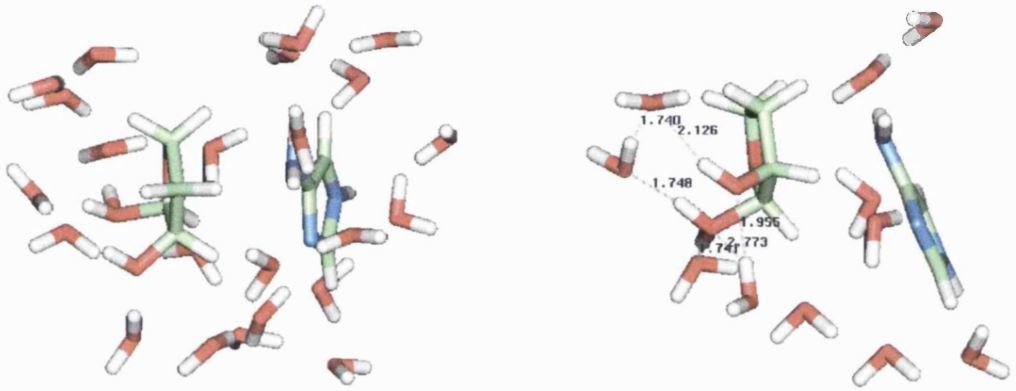


Figure 11.10: The solvated protonated alcohol and leaving group structure at AM1 QM/MM.

The solvated glycosyl cation possesses a series of tight 'water chains' between the hydroxyl and protonated hydroxyl groups, Figure 11.10. The nitrogen groups on the aglycone leaving group all possess strong hydrogen bonds.

11.6 Conclusion

In line with the study by Barnes and Williams', the GRACE/CHARMM calculated kinetic isotope effects indicated that the stepwise resolved mechanism was the preferred mechanism. The difference between the Barnes and Williams' and GRACE calculated isotope effects, relating to slightly different transition-states, lends credence to the existence of a 'family' or group of similar transition-states for a single mechanisms which is dependant on the solvent structure. The bonding to the leaving group in the stepwise mechanism is much looser than that found in the concerted mechanism, whereas the bond to the nucleophile is much shorter. This causes the considerable oxycarbenium character

for both the stepwise and concerted mechanisms in the ribose ring at transition-state which is manifests itself in the value for the secondary β-deuterium effect. The value for the concerted mechanism's secondary α-tritium effect and secondary α-deuterium is greater than that for the stepwise mechanism. The calculated isotope effects for the mechanisms, questions the validity of assigning mechanistic categories to reactions solely on their kinetic isotope effects. The model chosen to mimic the glycoside hydrolysis reaction totally neglects the possible effect that the phosphate group can have on the reaction.[121] The effect of the phosphate group may be to destabilise or stabilise one of the pathways so that only one mechanism is viable. A study involving the 'whole' system, Figure 11.2, would attempt to address the problem of whether the phosphate group has an effect or not.

Chapter Twelve

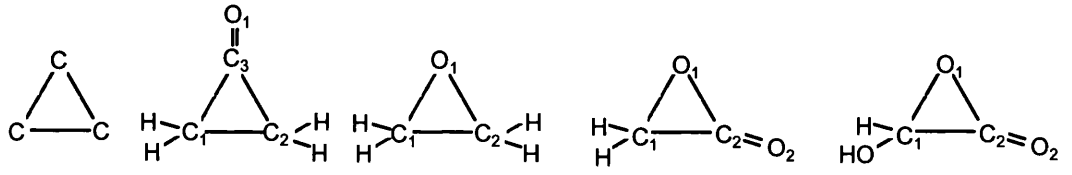
Electronic Insight into Three-membered Rings

Chapter 12 - Electronic Insight into Three-membered Rings

12.1 Introduction

The introduction of an exocyclic double bond onto a three-membered ring such as oxirane leads to compounds with an even larger ring-strain energy.[122] Oxiranones (α-lactones) fall into this category of compounds, being very unstable, highly reactive intermediates, and therefore difficult to isolate. This instability is due to the possibility of facile ring opening to a dipolar, zwitterionic species. The intermediacy of an α-lactone in halide elimination from α-halocarboxylates has been found to be dependent on the substituents attached to the α-carbon and on the properties of the solvent.[123] This intermediate has been described as a zwitterion,[124] an α-lactone[125] or an α-lactone possessing much ionic character.[126] MP2=full/6-311++G(d,p) calculations, both in the gas-phase and in solution using single-point calculations at IPCM by Firth-Clark *et al.*[127] have suggested that the intermediate is α-lactone in character. Firth-Clark *et al.* used hydroxyoxiranone as a model for a possible intermediate which might be involved in the heterolysis of glycosides of N-acetyl-α-D-neuraminic acid. They concluded that hydroxyoxiranone is an α-lactone with a long C-O bond length of 1.568Å for the *anti*-hydroxyoxiranone isomer.

Interest in the reactivities of small three-membered rings has led to the calculation of the standard enthalpies of formation of oxirane, oxiranone, cyclopropanone and hydroxyoxiranone at QCISD(t)=full/6-311++G(2df,p)//MP2=full/6-311++G(d,p) to an accuracy of ±10 kJ mol^{-1}.[128] Substituting these values into appropriate isodesmic equations allowed the calculation of the strain energies, kJ mol^{-1}, relative to oxirane. The results indicate hydroxyoxiranone (104) < cyclopropane (114) < oxirane (115) < oxiranone (169) < cyclopropanone (181). It appears that, hydroxyoxiranone has a significantly low strain energy as compared to oxiranone.



Cyclopropane Cyclopropanone Oxirane Oxiranone Hydroxyoxiranone

Figure 12.1: Three-membered rings studied.

The structural properties of these compounds are a very interesting topic which can now be addressed using Bader's 'atoms in molecules' theory to compute the electronic topological characteristics.[129] The investigation of the role of solvent is carried out using a self-consistent isodensity polarised continuum model. By this means it is possible to rationalise the structure and bonding in a series of three-membered rings, Figure 12.1, and offer an explanation to hydroxyoxiranone's curiously low strain energy.

12.2 Methods

All calculations were carried out using Gaussian 94.[18] Gas-phase geometry optimisations were performed without any symmetry constraints at the Hartree-Fock level using the 6-31+G(d,p) basis set with MP2 electron correlation. The effects of aqueous solvation upon structure were determined by means of geometry optimisation using a self-consistent isodensity polarised continuum model (SCI-PCM) at the Hartree-Fock level using the 6-31+G(d,p) basis set with a dielectric constant of 78.36 for water. Single-point energies were then calculated at MP2/6-31+G(d,p). The self-consistent reaction field method models the solvent as a continuum of uniform dielectric constant. The SCI-PCM method includes the effects of solvation in the solution to the self-consistent field problem. The solute charge distribution is allowed to be polarised by the reaction field and the calculation is iterated until self-consistency is achieved. The electron topology was determined using the AIMPAC series of programs which invokes Bader's 'theory of atoms in molecules'.

12.3 Atoms in molecules analysis

Matter is composed of atoms. This is a consequence of the manner in which electrons are distributed throughout space in an attractive field exerted by the nuclei. These nuclei act as point attractors immersed in a cloud of negative charge, the electron density. This electron density describes the manner in which the electronic charge is distributed throughout real space determining the appearance and form of matter. The structure and bonding of molecules can best be described in terms of topological characteristics of the electron density (ρ), its Laplacian ($\nabla^2\rho$) and the ellipticity (ε) of the molecules bonds.

12.3.1 Electron density, ρ

The electron density, $\rho(r)$, is a physical quantity having a definite value at each point in space. It is a scalar field over three dimensional space. The topological properties of this scalar field are conveniently summarised in terms of the number and type of points called critical points. The structure of bonds is made more evident and amenable to analysis through the study of the associated gradient vector of the charge density, $\nabla\rho(r)$. A representation of 'lines of force' is a display of a vector field. The properties of a vector field are characterised by associating a direction as well as a magnitude with each point in space. The display of $\nabla^2\rho(r)$ for a molecule makes visible the definitions of its atoms and of a particular set of lines linking certain pairs of nuclei within the molecule.

12.3.2 Critical points

There are four possible kinds of stable critical points in $\rho(r)$, each being associated with an important topological element of structure. The distribution of charge within a molecular system is an indication of the forces acting within the system. Dominant among these is the attractive force exerted by the nuclei, a consequence of the localised nature of the nuclear charge. This interaction is responsible for the most important topological property exhibited by the molecular charge distribution in a many electron system: that of $\rho(r)$ exhibiting a local maximum only at positions of the nuclei. Whether it is a maximum, minimum or a saddle point, each of the topological features of $\rho(r)$ has associated with it a point in space called a critical point. At this point, denoted by the position vector, r_c, the first derivatives of $\rho(r)$ i.e. $\nabla\rho(r) = 0$, where $\nabla\rho$ denotes the operation in Equation 12.1. Whether a function is a maximum or minimum at an extremum is determined by the sign of its second derivative.

$$\nabla\rho = \mathbf{i}\,\delta\rho/\delta x + \mathbf{j}\,\delta\rho/\delta y + \mathbf{k}\,\delta\rho/\delta z \qquad \text{Equation 12.1}$$

A bond critical point can be found between every pair of nuclei which are considered to be linked by a chemical bond. Here the electron density is at a maximum at position vector, r_c, and a minimum perpendicular to the bond. A ring critical point is formed when nuclei are linked together to form a ring. Here the charge density at this point is at a minimum at r_c, and a maximum perpendicular to the plane. An example of this can be seen in diborane, B_2H_6, Figure 12.2.

9.2.1 GRACE

GRACE is based on a mathematical program whose functions have been oriented to the needs of QM/MM modelling.[101] The main interface of GRACE is via a tool control language which is a structured programming language. It provides a sophisticated 'nerve centre' which uses external codes as well as its own internal algorithms. Visualisation codes are not part of GRACE; however, it can interface with programs such as XMol[102] or RasMol[103] via a series of pdb files which can be read and animated. This makes the visualisation of eigenvectors during saddle point location possible ensuring that the correct mode is followed. The interface with GAMESS-UK[104] operates as part of the *ab initio* QM/MM modelling method in GRACE. When the energy and gradient of a QM/MM system is required, GRACE constructs an appropriate input deck for GAMESS-UK, runs the code and then parses out the energy and gradients from the log file. The molecular mechanics components of the QM/MM function are then constructed via an interface with CHARMM. (See Appendix B for more details on GRACE)

9.2.2 AM1/COSMO

A reaction profile of the S_N2 nucleophilic substitution reaction is computed using the AM1/COSMO method and the approximate saddle point geometry is used initiate a saddle search. Once a transition-state is located, a set of 'valence' internal co-ordinates is then generated using GRACE and CAMVIB which remove the rotational and translational contamination from the hessian before the kinetic isotope effects are computed by CAMISO.

9.2.3 AM1 QM/MM

One approach to model a reaction in solution is to use a combined quantum mechanical (QM) and molecular mechanical (MM) method, Figure 9.1. This 'water droplet' method is split into two regions. One region, the reacting core, contains all the reacting atoms of interest. The atoms in the reacting core are subjected to the quantum mechanical simulation. The surrounding region contains all the atoms outside the reacting core. In this case they are solvent molecules. A convenient way to describe the behaviour of a system being studied with a hybrid potential, as in CHARMM, is to use a notation involving an effective Hamiltonian, H_{eff}.[105] In this formulation, the energy of the

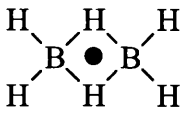


Figure 12.2: Structure of diborane showing a ring critical point denoted by •.

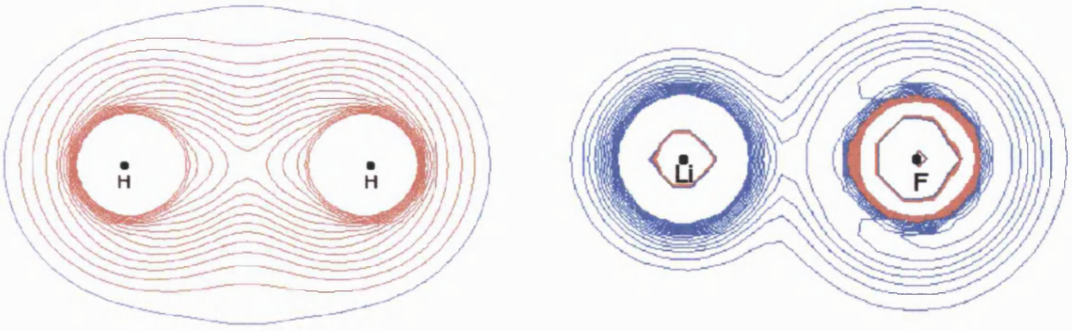
A cage critical point can be found when the charge density is a local minimum at the centre of the cage structure as in tetrahedrane, C_4H_4. Finally, there is a critical point located over an atom which is a maximum of electron density. Here the Coulombic potential becomes infinitely negative when an electron and nucleus coalesce (i.e. this is the electron density around an atoms nucleus).

12.3.3 The Laplacian

The value of the quantity $\nabla^2\rho$, called the Laplacian of ρ in Equation 12.2 is denoted by:

$$\nabla^2\rho = \nabla \bullet \nabla\rho = \delta\rho^2/\delta x^2 + \delta\rho^2/\delta y^2 + \delta\rho^2/\delta z^2 \qquad \text{Equation 12.2}$$

The Laplacian ($\nabla^2\rho$) of the electron density recovers the shell structure of an atom by displaying a corresponding number of alternating shells of charge concentration and charge depletion. The uniform sphere of charge concentration present in the valence shell of a free atom is distorted upon chemical combination to form local maxima and minima. This scalar function, $\nabla^2\rho$, measures the local concentration of charge. A negative value of the function at some point implies a higher value of the density compared to the immediate surroundings, and *vice versa* for positive values of $\nabla^2\rho$.[130] A covalent bond implies that there is a build-up of charge at the bond critical point, and therefore it will have a 'large' negative value of $\nabla^2\rho$, whereas a large positive value at the bond critical point would be a characteristic of an ionic bond, a hydrogen bond or a van der Waals molecule.[131] This can be seen in the Laplacian diagrams for a covalent hydrogen molecule ($\nabla^2\rho$ = -1.24) and the ionic interaction between lithium fluoride ($\nabla^2\rho$ = +0.63). The Laplacian contour diagram shows that for the hydrogen molecule the electron density is equally spread between the hydrogen atoms whereas in lithium fluoride, the electron density is distributed towards the more electronegative fluorine atom, Figure 12.3.



Laplacian for the Hydrogen Molecule Laplacian for Lithium Fluoride

Figure 12.3: *In vacuo* contour line diagrams of the calculated [MP2/6-31+G(d,p)] Laplacian concentration. The red and blue lines are regions in which electron charge is concentrated or depleted, respectively.

12.3.4 Ellipticity, ε

The ellipticity, ε, measures the asymmetry between the two principal curvatures (λ_1, λ_2) of electron density perpendicular to the bond and is defined as $\varepsilon = \lambda_1/\lambda_2 - 1$. In a single bond, such as the C-C bond in ethane, $\varepsilon = 0$, as there is no asymmetry of electron density perpendicular to the plane. Higher values of ε can be found for the C=C double bond in ethene $\varepsilon = 0.41$, as this ethene bond possess π-electron density perpendicular to the plane. Therefore ε can be used as a measure of π-character of a bond.[132]

12.3.5 Bond path and angle deviation

It is possible to define the lines of maximum electron density between two nuclei, usually called bond paths. The bond path, r_b, is defined as the line of maximum electron density between two nuclei. This will not usually coincide with the geometrical bond path, r_e, which is a straight line drawn between the nuclei.[133] The bond angle path, α_b, is calculated by the angle generated between tangents to the bond path at some nucleus. The difference between the geometrical bond angle, α_a, at some nucleus and the bond path angle, α_b, may be used to quantify strain, Δ_α, Figure 12.4.

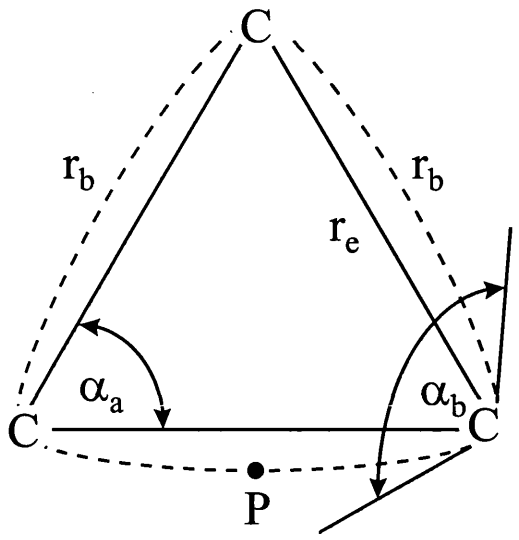


Figure 12.4: Geometrical and bond parameters of a typical three-membered ring. The bond paths are indicated by the dashed lines and α_a is the geometric angle, α_b is the bond path angle, r_e is the geometric distance, r_b is the bond path length and P is the bond critical point.

12.3.6 Bader Charges

Integrated properties were calculated using PROAIMV, a subprogram of the AIMPAC suite. The boundaries of an atomic subsystem (Ω) are defined such that the subsystem obeys the 'zero-flux' condition, Equation 12.3, for all points r on the surface (**n** is a vector normal to this surface).

$$\nabla\rho(\mathrm{r})\bullet\mathbf{n}(\mathrm{r}) = 0 \qquad \text{Equation 12.3}$$

Sub-systems so defined obey the Virial theorem. A property density, $\rho_A(\mathrm{r})$, corresponding to an observable, $\hat{A}$, can be integrated over the basin defined by the zero-flux surface to yield the atomic expectation value of $\hat{A}$, Equation 12.4. In this fashion, it is possible to evaluate atomic properties such as populations and charges.

$$A(\Omega) = \int_{\Omega} \mathrm{d}\tau\rho A(\mathrm{r}) \qquad \text{Equation 12.4}$$

In this chapter we focus on the net charge on an atom which is given by the sum of the nuclear charge and the electron population of the atom. For example, if the atomic basin of oxygen contains 8.5 electrons, then its net charge is -0.5.

12.4 Results

The effect of solvation upon several structures was determined by means of geometry optimisation using a self-consistent isodensity polarised continuum model. The energies of all the *in vacuo* gas and *in aquo* structures are given in Table 12.1. Table 12.2 contains the ring critical points (RCP), bond lengths, bond path lengths electron density (ρ), Laplacian and ellipticity (ε) of the studied three-membered rings for the carbon – carbon and carbon – oxygen bonds. Table 12.3 contains the bond path angles (Δ_α) for the series while Table 12.4 contains the Bader and Mulliken charges for the carbon and oxygen atoms.

Table 12.1: Calculated energies for the studied series

	In vacuo		*In aquo*	
	HF/6-31+G(d,p)	MP2/6-31+G(d,p)	HF/6-31+G(d,p)	MP2/6-31+G(d,p)b
Cyclopropane	-117.071002	-117.503207	-117.072144	-117.503993
Cyclopropanone	-190.735657	-191.325978	-190.746440	-191.329836
Oxirane	-152.876471	-153.346287	-152.884272	-153.349314
Oxiranone	-226.570261	-227.205337	-226.587392	-227.208238
anti-Hydroxyoxiranone	-301.434385	-302.254943		
gauche-Hydroxyoxiranone	-301.433783	-302.252983	-301.460374	-302.261256

[a] energies in hartrees = 1 hartree = 2625 kJ mol-1,[b] single point energy.

Table 12.2: Critical point analysis for carbon – carbon and carbon – oxygen bonds.

		In vacuo [MP2/6-31+G(d,p)]						*In aquo* [MP2/6-31+G(d,p)]					
	Bond	RCP $\rho(r)$	Bond Length /Å	Bond Path Length /Å	$\rho(r)$	$\nabla^2\rho(r)$	ε	RCP $\rho(r)$	Bond Length / Å	Bond Path Length /Å	$\rho(r)$	$\nabla^2\rho(r)$	ε
Cyclopropane	$C_1 - C_2$	0.198	1.507	1.514	0.238	-0.431	0.450	0.204	1.500	1.508	0.247	-0.523	0.500
Cyclopropanone	$C_1 - C_2$	0.190	1.471	1.484	0.265	-0.582	0.223	0.199	1.456	1.472	0.282	-0.742	0.227
	$C_1 - C_3$		1.577	1.581	0.209	-0.294	0.621		1.561	1.567	0.221	-0.391	0.580
	$C_3 - O_1$		1.218	1.218	0.395	+0.338	0.048		1.188	1.188	0.282	+0.742	0.227
Oxirane	$C_1 - C_2$	0.209	1.468	1.484	0.256	-0.576	0.255	0.224	1.451	1.473	0.273	-0.689	0.268
	$C_1 - O_1$		1.446	1.449	0.244	-0.388	0.710		1.412	1.415	0.254	-0.433	1.208
Oxiranone	$C_1 - C_2$	0.197	1.450	1.471	0.275	-0.660	0.108	NCP	1.429	1.456	0.291	-0.837	0.115
	$C_1 - O_1$		1.567	1.624	0.197	+0.107	6.592		1.501	NCP	NCP	NCP	NCP
	$C_2 - O_1$		1.346	1.352	0.303	-0.364	0.228		1.296	1.299	0.337	+0.120	0.356
	$C_2 - O_2$		1.204	1.204	0.410	+0.268	0.102		1.178	1.179	0.443	+0.495	0.075
anti-Hydroxyoxiranone	$C_1 - C_2$	NCP	1.454	1.473	0.280	-0.692	0.103						
	$C_1 - O_1$		1.582	NCP	NCP	NCP	NCP						
	$C_2 - O_1$		1.344	1.352	0.304	-0.366	0.152						
	$C_2 - O_2$		1.207	1.207	0.407	+0.241	0.104						
gauche-Hydroxyoxiranone	$C_1 - C_2$	NCP	1.450	1.470	0.282	-0.712	0.117	NCP	1.433	1.462	0.301	-0.897	0.132
	$C_1 - O_1$		1.572	NCP	NCP	NCP	NCP		1.519	NCP	NCP	NCP	NCP
	$C_2 - O_1$		1.346	1.349	0.302	-0.362	0.177		1.294	1.296	0.339	+0.126	0.294
	$C_2 - O_2$		1.205	1.205	0.409	+0.265	0.108		1.179	1.179	0.442	+0.493	0.076

$\rho(r) = e / a_0^3 = 1.081 \times 10^{12}$ Cm^{-3}, $\nabla^2\rho(r) = e / a_0^5 = 3.8611 \times 10^{32}$ cm^{-5}, ε is dimensionless, NCP = No Critical Point Detected, RCP = Ring Critical Point.

Table 12.3: Bond path angles (Δ_α) for the cyclic series.

		In vacuo	*In aquo*
	Angle	Δ_α (deg)	Δ_α (deg)
Cyclopropane	C – C – C	16.1	18.2
Cyclopropanone	$C_1 - C_2 - C_3$	20.5	23.0
	$C_1 - C_3 - C_2$	18.4	20.8
Oxirane	$C_1 - C_2 - O_1$	10.3	11.1
	$C_1 - O_1 - C_2$	13.2	13.1
Oxiranone	$C_1 - C_2 - O_1$	10.0	14.5
	$C_2 - C_1 - O_1$	-19.3	
	$C_1 - O_1 - C_2$	2.5	
anti-Hydroxyoxiranone	$C_1 - C_2 - O_1$	11.6	
gauche-Hydroxyoxiranone	$C_1 - C_2 - O_1$	10.6	14.6

Table 12.4: Bader and Mulliken charges.

		In vacuo		*In aquo*	
	Atom	Bader Charges	Mulliken Charges	Bader Charges	Mulliken Charges
Cyclopropane	C	+0.02	-0.29	+0.09	-0.29
Cyclopropanone	C_1	-0.03	-0.21	+0.06	+0.15
	C_3	+1.04	+0.11	+1.18	+0.28
	O_1	-1.11	-0.39	-1.37	-0.58
Oxirane	C_1	+0.43	-0.09	+0.62	-0.05
	O_1	-0.91	-0.39	-1.21	-0.51
Oxiranone	C_1	+0.29	-0.06	+0.47	+0.03
	C_2	+1.65	+0.46	+2.00	+0.61
	O_1	-0.91	-0.33	-1.23	-0.46
	O_2	-1.15	-0.43	-1.40	-0.62
anti-Hydroxyoxiranone	C_1	+0.87	+0.24		
	C_2	+1.64	+0.44		
	O_1	-0.91	-0.34		
	O_2	-1.15	-0.43		
gauche-Hydroxyoxiranone	C_1	+0.87	+0.16	+1.15	+0.36
	C_2	+1.67	+0.54	+2.02	+0.69
	O_1	-0.91	-0.33	-1.24	-0.49
	O_2	-1.15	-0.43	-1.40	-0.62

12.5 Discussion

To obtain a point of reference for values using Bader's theory, a series of simple, non-cyclic molecules were studied at MP2/6-31+G(d,p), Table 12.5.

Table 12.5: Table of values for simple molecules at MP2/6-31+G(d,p).

Molecule	Bond	$\nabla^2\rho(r)$	ε	Bader Charges			
				C			
CH_3CH_3	C – C	-0.669	0.00	+0.23			
$CH_2{=}CH_2$	C = C	-1.108	0.41	+0.05			
				C	O_1	O_2	
CH_3OH	C – O_1	-0.261	0.01	+0.78	-1.24		
HCHO	C = O_1	+0.364	0.08	+1.26	-1.26		
HCOOH	C – O_1	-0.270	0.11	+1.93	-1.30		
	C = O_2	+0.115	0.11			-1.35	
				C_1	C_2	O_1	O_2
CH_3CH_2OH	C_1 – C_2	-0.735	0.04	+0.21	+0.73		
	C_2 – O_1	-0.248	0.02			-1.24	
CH_3CHO	C_1 – C_2	-0.802	0.04	+0.18	+1.19		
	C_2 = O_1	+0.298	0.07			-1.30	
CH_3COOH	C_1 – C_2	-0.788	0.05	+0.21	+1.82		
	C_2 – O_1	-0.335	0.12			-1.28	
	C_2 = O_2	+0.111	0.13				-1.35

The ellipticity for a C-C single bond in ethane has a value of 0 while the C=C double bond in ethene has a value of 0.41 This is due to π-character of the double bond in ethene. The values for a C-O single bond are close to that for a C-C single bond. When the C=O and C-O bond are attached to the same carbon, the value of ε for both bonds increases. The analysis of simple molecules shows that the Laplacian for the carbonyl group (C=O) is positive whereas for the carbon-oxygen single bond (C-O) it is negative. The Bader charges on the oxygen in C=O or C-O bonds are approximately equal. The negative charge of the oxygen is counterbalanced by the positive charge on the carbon atom. If the carbon is attached to a methyl group some of this positive charge is passed onto the carbon of the methyl group.

12.5.1 Atoms in molecules analysis

The moleculargraphs of cyclopropane, cyclopropanone, oxirane, oxiranone and hydroxyoxiranone were obtained from MP2 electron densities. Bond path lengths are in Ångströms and bond path angles are in degrees. The geometrical bond lengths and bond angles are given within parentheses. (●) denotes the electron densities (e a.u.$^{-3}$) at the bond critical point while (▲) denotes a ring critical point, (●) denotes a carbon atom, (O) denotes a oxygen atom and the ellipticities are given in square brackets. The hydrogen atoms are omitted for simplicity. In the Laplacian contour diagrams, the red and the blue lines are regions in which electron charge is concentrated or depleted, respectively.

12.5.1.1 Cyclopropane

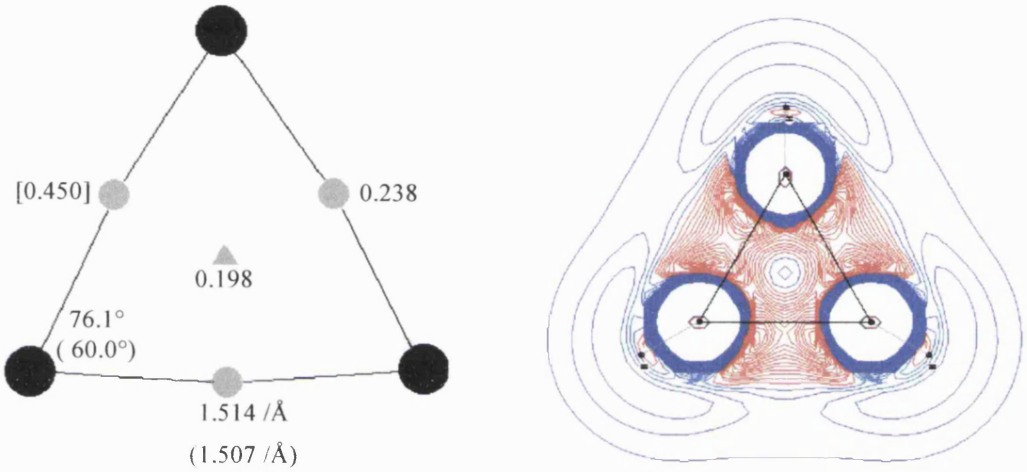


Figure 12.5: *In vacuo* moleculargraph and contour line diagram of the calculated Laplacian concentrations for cyclopropane.

The prototype system for a strained three-membered ring is considered to be cyclopropane. The C-C bond path lengths (1.514Å) are slightly longer than the interatomic distances (1.507Å). The strain of the system is evident in the bond path angles which are larger than the conventional bond angles by 16.1°. Although highly strained, with a strain energy of 114 kJ mol^{-1}, topologically it possesses a ring critical point with a density of 0.198. The density of the bond critical points are larger and positioned away from the ring critical point creating a stable, closed ring system. The bond ellipticities of 0.450 are slightly greater than ethene indicating a significant π-character in C-C the bonding.

Contour line diagrams of the calculated Laplacian concentrations show spheres of charge concentration and depletion surrounding each atom in a molecule. These spheres have been associated with inner core and outer valence electrons. The spheres are distorted in ways which provide insight into electronic and bonding patterns.[134] If the nuclei are linked to form a ring, then a ring critical point is found in its interior. From the gas-phase diagram of the Laplacian, a ring critical point can be found at the centre of cyclopropane, Figure 12.5. The Laplacian for C-C bonds are all negative showing the covalent character of the bonding. The effect of solvation on the system is to increase the electron density at the critical points which is then reflected in the increased value of their bond ellipticities. This increase in the electron density causes the bonds to bow outwards in the plane of the bonds, increasing the bond path angles and hence strain of the system, Figure 12.6.

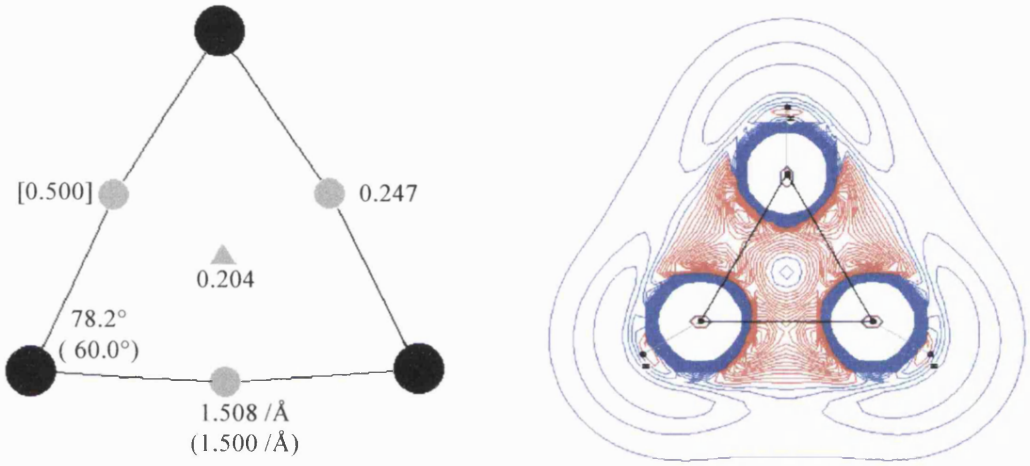


Figure 12.6: *In aquo* moleculargraph and contour line diagram of the calculated Laplacian concentrations for cyclopropane.

12.5.1.2 Cyclopropanone

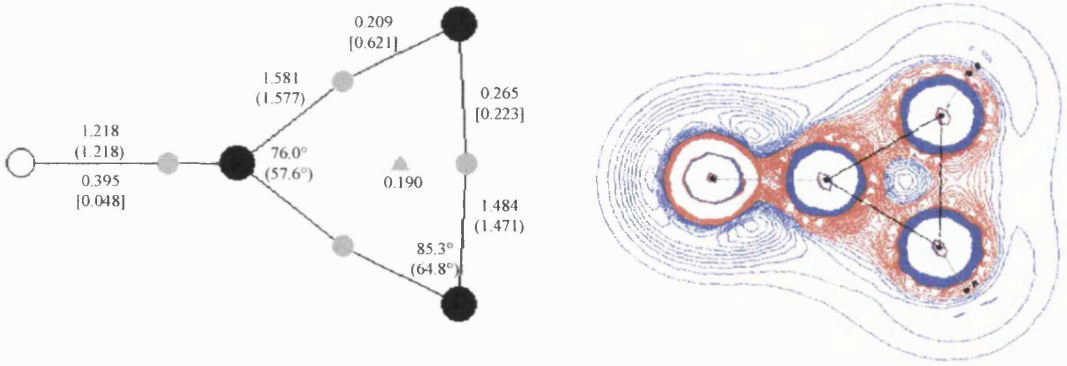


Figure 12.7: *In vacuo* moleculargraph and contour line diagram of the calculated Laplacian concentrations for cyclopropanone.

In comparison to cyclopropane, there is now a significant decrease in the difference between the density of the ring critical point, 0.190, and that of the apex C-C bond critical point, 0.209. The base and apex bond path lengths of the C-C bonds are all greater than the bond lengths, producing a larger curvature in the bond path angle 85.3° and in the maximum electron density path. The ellipticity in the apex C-C bond has also increased to 0.621, Figure 12.7. This produces an overall more strained system than cyclopropane. The Laplacian for the carbonyl group (C=O) is positive implying a resonance, ionic character of the bond which is reflected in the Bader charges on carbon (+1.04) and oxygen (-1.11) atoms. This is a standard position for all carbonyl groups. The effect of solvation on the system is to increase the electron density at the critical points of all the bonds. This is reflected in the increase of the value for bond ellipticities, Figure 12.8.

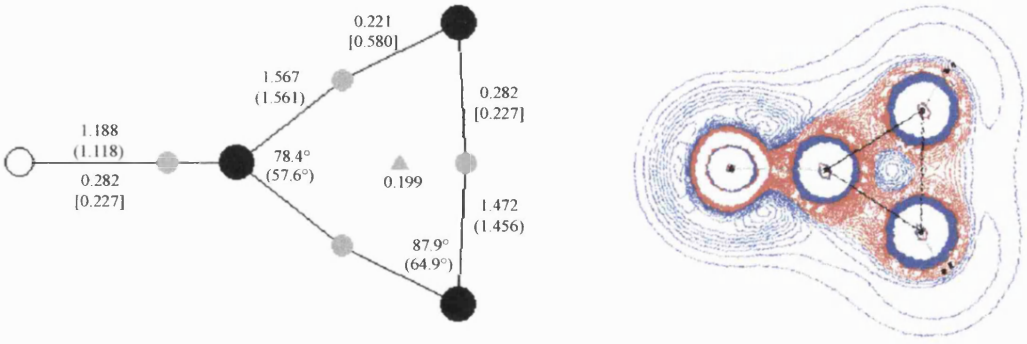


Figure 12.8: *In aquo* moleculargraph and contour line diagram of the calculated Laplacian concentrations for cyclopropanone.

12.5.1.3 Oxirane

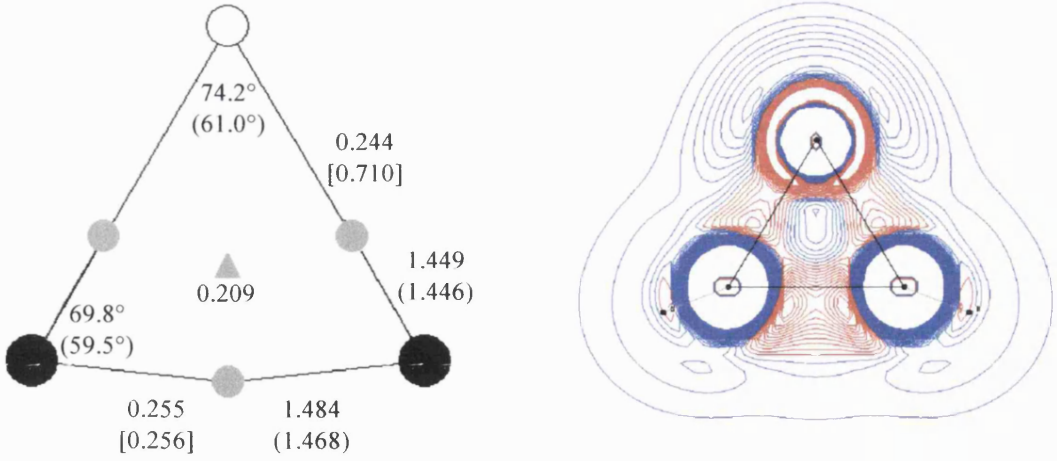


Figure 12.9: *In vacuo* moleculargraph and contour line diagram of the calculated Laplacian concentrations for oxirane.

Oxirane is isoelectronic with cyclopropane possessing a similar strain energy. The bond path length of the C-C bond is slightly elongated. Like cyclopropane, there is a significant difference between the density in the ring critical point, 0.209, and that at the C-O bond critical point, 0.244, or at the C-C bond critical point, 0.256. The difference in the bond path angles in oxirane, CCO (10.3°) and COC (13.2°) are less than in cyclopropane. The Laplacian for all the bonds are negative showing the covalent character of the bonding, Figure 12.9. The ellipticity of the C-O bond (0.710) as compared to the C-C (0.256) shows that the π-electron density is preferentially accumulated along the C-O bond axis. The effect of solvation on the system is to increase the electron density of all the bond critical points and at its ring critical point, Figure 12.10, and the ellipticity if the C-O bonds.

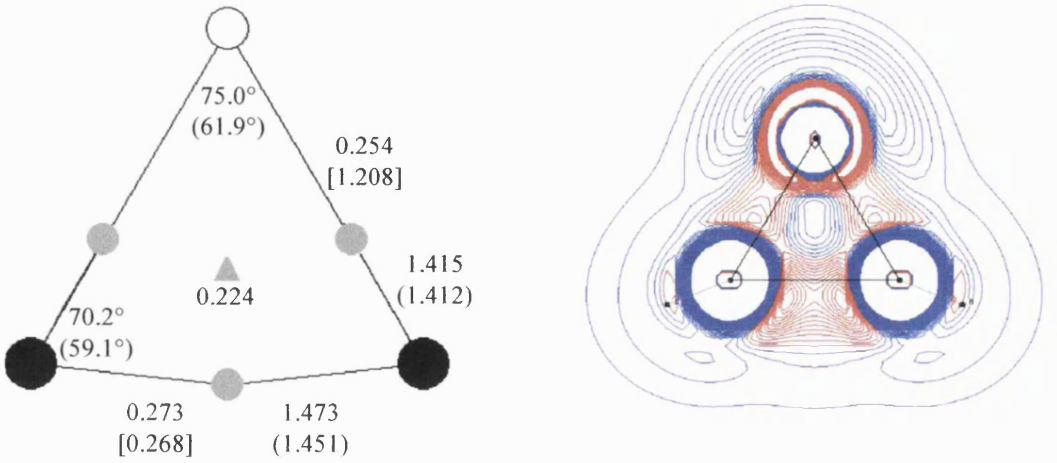


Figure 12.10: *In aquo* moleculargraph and contour line diagram of the calculated

Laplacian concentrations for oxirane.

12.5.1.4 Oxiranone

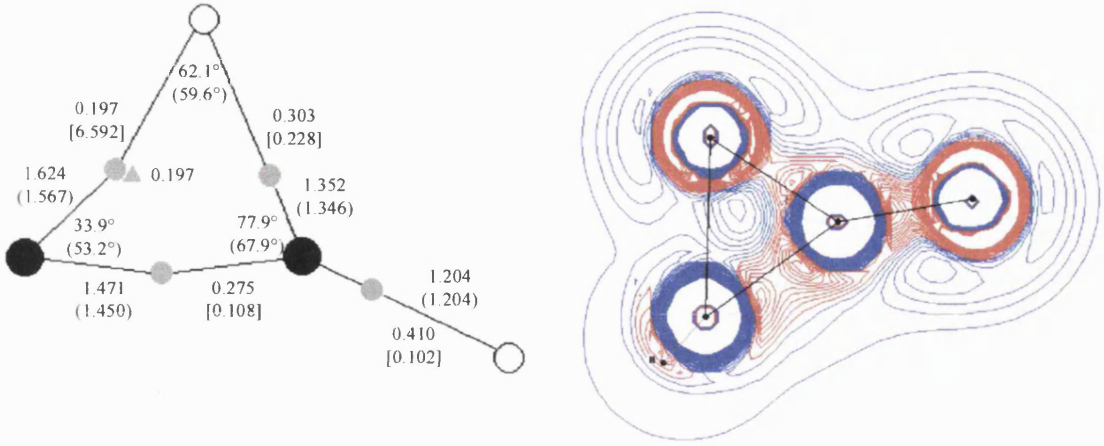


Figure 12.11: *In vacuo* moleculargraph and contour line diagram of the calculated Laplacian concentrations for oxiranone.

The bonding in oxiranone is very interesting. The C_1-O_1 bond is of unusual length (1.567Å). However, comparing it to the even longer bond path length (1.624Å), there is a difference of 0.057Å. The opening of a ring structure is the result of the coalescing of a ring and bond critical point. The positive curvature of the ring annihilates the in-plane negative curvature of the bond point to yield a zero curvature characteristic of an unstable critical point. This decrease in magnitude of negative curvature and its eventual disappearance means that the ε of the bond which is to be broken increases dramatically becoming infinite at the geometry of the bifurcation point. Thus a structure possessing a bond with an unusually high ε is potentially unstable.[135] This can be seen for the C_1-O_1 bond in oxiranone which has a very large ε, (6.592), Figure 12.11. The value of the electron density at the C_1-O_1 bond critical point and oxiranones ring critical point only differ in energy by 0.0002 e $a.u^{-3}$. This implies a nearly flat bottomed trough in the distribution of electron density linking these two critical points. Of the systems studied, oxiranone is the only one that has a negative bond path angle (-19.3°). This infers that the electron density of the bond bows inwards. The C_1-O_1 bond critical point possesses an appreciable electron density and a positive Laplacian which is indicative of a closed shell, ionic interaction. Oxiranone in the gas-phase can be best described as a closed ring, α-lactone which possesses an ionic character between the C_1 and O_1 atoms.

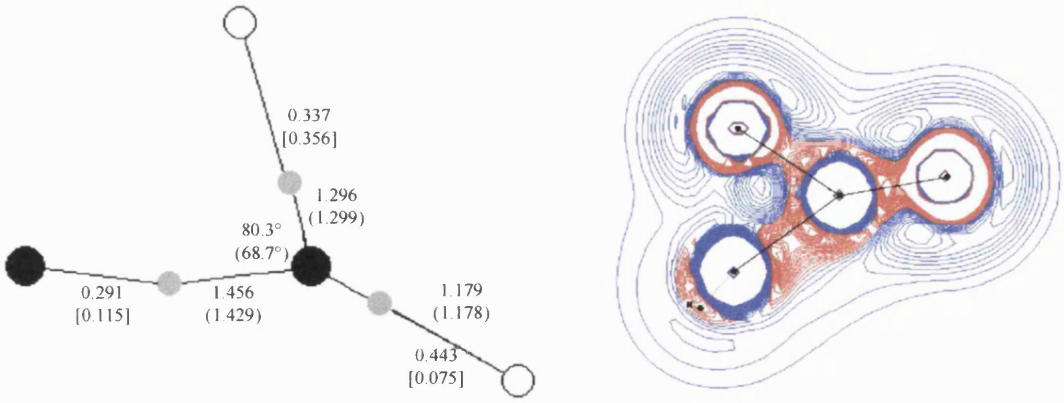
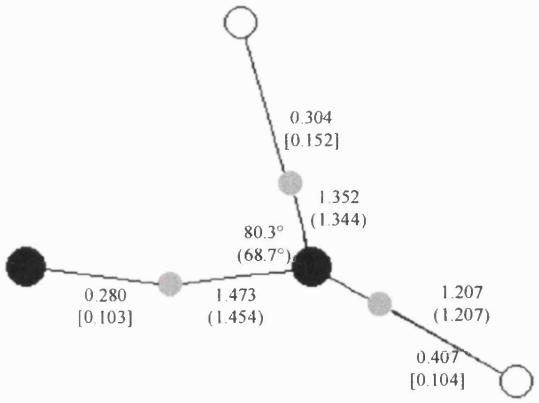


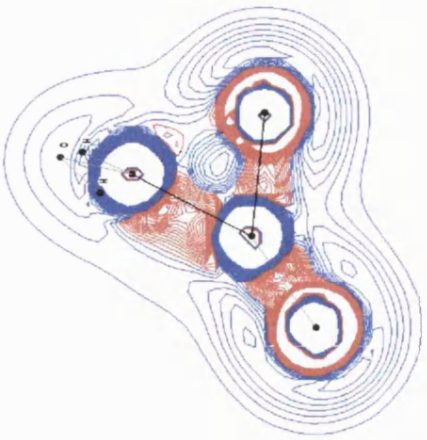
Figure 12.12: *In aquo* moleculargraph and contour line diagram of the calculated Laplacian concentrations for oxiranone.

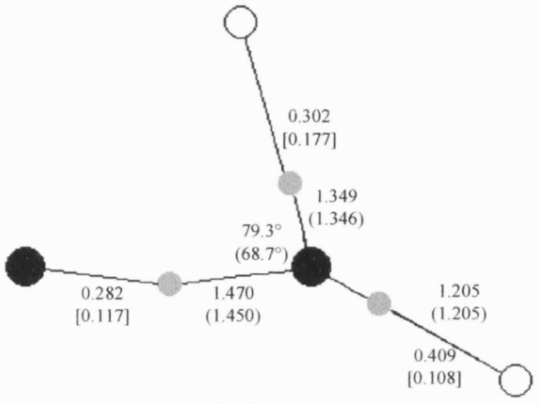
In solution, the geometrical bond length of the C_1-O_1 bond is 1.501Å. This is a reasonable distance for a C-O bond in a cyclic system. However, in solution there is no critical point found between these atoms. The inclusion of solvent effects causes the migration of the ring point along the trough to coalesce with the bond critical point, yielding an open structure, Figure 12.12. This open structure can best be described as a ring-opened, zwitterion, as there is no bonding interaction between the C_1 and O_1 atoms. The Laplacian of the C_2-O_1 bond changes from negative to positive indicating a double bond type structure and resonance effect for the carboxylate moiety.

12.5.1.5 Hydroxyoxiranone



anti-Hydroxyoxiranone





gauche-Hydroxyoxiranone

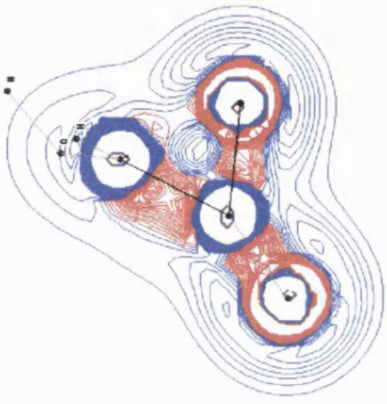
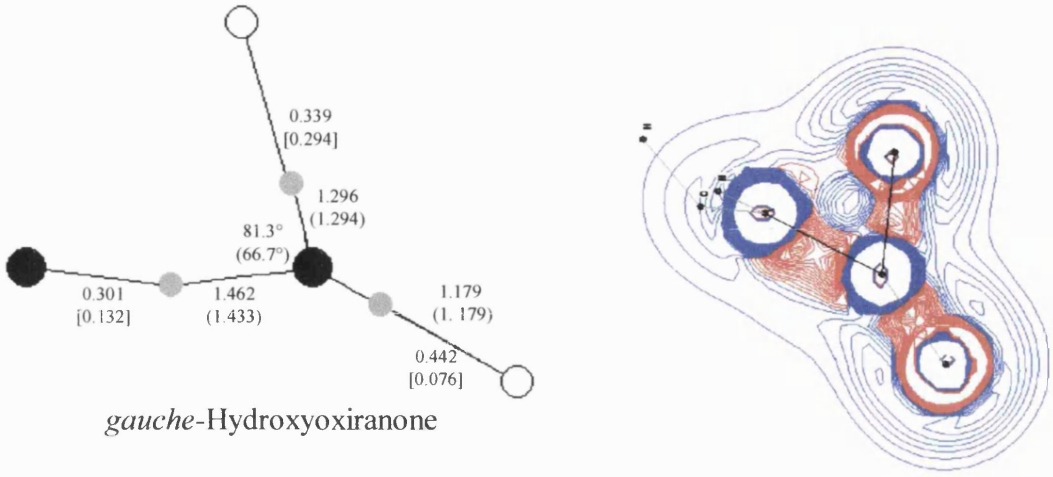


Figure 12.13 *In vacuo* moleculargraph and contour line diagram of the calculated Laplacian concentrations for hydroxyoxiranone.

The *anti*-isomer has slightly lower energy than the *gauche*-isomer by 5 kJ mol^{-1}. The surface delocalisation of σ-electrons leads to the enhancement of stabilisation in three-membered rings. Substituents with σ-electron-withdrawing ability remove electrons from the surface orbital of ring system thus destabilising the ring. Therefore substituents with σ-electron-donating ability push electrons into the ring system increasing ring stability. The electron-donating hydroxy substituent in hydroxyoxiranone preferentially stabilises the C_1 carbon by delocalising the positive charge to such an extent that no C_1-O_1 interaction exists. This leads to an open zwitterionic and rather than a cyclic α-lactone structure for hydroxyoxiranone, Figure 12.13. Three-membered rings have a great deal of angle strain, since 60° angles represent a large departure from the tetrahedral angle of 109.5°. The ring opening of a three-membered ring such as hydroxyoxiranone will relieve this strain giving it significantly lower strain energy. The effect of solvation on the system is to increase the overall zwitterionic character of the

system, Figure 12.14. This can be seen in the positive Laplacian for both C-O bonds showing a delocalisation of electron density over both oxygens and an increase in the carbonyl nature of the bonding.



gauche-Hydroxyoxiranone

Figure 12.14 *In aquo* moleculargraph and contour line diagram of the calculated Laplacian concentrations for hydroxyoxiranone.

12.6 Conclusion

Baders 'theory of atoms' in molecules shows that cyclopropane, cyclopropanone and oxirane are all closed-ring systems in the gas-phase, whereas hydroxyoxiranone is an open ring system and should be classed as a zwitterion. Oxiranone is an intermediate case where the Bader analysis of the bonding in the gas-phase suggests that it is an α-lactone possessing much ionic character. In solution, cyclopropane, cyclopropanone and oxirane are still all closed ring systems, whereas oxiranone and hydroxyoxiranone are open ring systems that should be classed as zwitterions and not α-lactones. The opening of hydroxyoxiranone's ring structure relieves the ring-strain in this system and this gives rise to its significantly lower strain energy. As previously stated, oxiranone and substituted oxiranones have been described as zwitterionic, as a α-lactone or as a α-lactone possessing much ionic character. Each of the above descriptions are valid for certain systems. If the substituent is electron-donating such as NH_2 then this stabilises the zwitterionic over the α-lactone form. Therefore the compound is best described as zwitterion. If the substituent is electron-withdrawing such as NO_2 then this destabilises the zwitterionic over the α-lactone form. Therefore the compound is best described as an α-lactone. In changing from the zwitterionic over to the α-lactone form, there must be a

transitional point which is best described as an α-lactone possessing much ionic character. If the strain energy of nitro (α-lactone) and amino-oxiranone (zwitterion) where calculated, then the nitro would have a greater strain energy while the amino would be less stained, relative to oxiranone.

Chapter Thirteen

Conclusions and Future Work

Chapter 13- Conclusions and Future Work

13.1 Attainment of Goals

The aim of this thesis was to show that modern computational techniques can be used to help understand and interpret fundamental chemical processes in solution using a mixture of *ab initio* and semi-empirical levels of theory along side continuum and hybrid quantum mechanical/molecular mechanical descriptions of aqueous solutions. This has been achieved, and the following objectives have been completed:

- examination the effect of solvation upon the mechanism of a reaction.
- interpretation the effects of secondary kinetic isotopes.
- assessment of structure-reactivity relationships.
- evaluation of transition-state models.

13.2 Nucleophilic substitution

Nucleophilic substitution of protonated alcohols have been studied using the AM1/COSMO method. The results showed the same qualitative trend as Ingold *et al*'s experimental solvolysis study of a series of alkylated bromides indicating the same mechanistic changeover between the ethyl and *iso*-propyl substrates. From the calculated results, it has been shown that small KIEs are indicative of tight transition-states and large KIEs correspond to loose transition-states, regardless of the mechanism. This questions the validity of assigning mechanistic categories to reactions solely on examining the KIEs.

13.3 Mechanistic changeover

Mechanistic changeover occurs at a point where barriers to two mechanisms are approximately equal and both mechanisms would therefore coexist with neither being in dominance. The AM1/COSMO energy surfaces for the identity reactions of protonated alcohols showed that two distinct, competing mechanistic pathways exist. These correspond to Ingold's S_N2 and S_N1 mechanistic pathways. The alternative mechanistic descriptions for changeover such as the Doering and Zeiss' 'merging' mechanistic pathway or a discontinuous jump from one mechanism to the other were not seen in any

of the energy surfaces. The effect of adding a methyl group to the reaction was too great a perturbation on the energy surface to allow this borderline to be studied in detail. Therefore, the identity reaction of substituted benzyl chlorides was used to investigate the precise point of mechanistic changeover. The methyl substituted benzyl chloride showed a borderline case where both S_N1 and S_N2 mechanisms run concurrently differing only slightly in activation energy. Before, at and past this transitional point, the surfaces showed a 'see-saw' effect between the S_N2 to S_N1 reaction mechanisms. None of the surfaces showed a single mechanism operating.

13.4 Structure-reactivity relationships

13.4.1 Hammett Correlations

It was demonstrated for the studied reaction, that Linear correlations in Hammett plots were shown to originate from a single mechanism operating. For solvolysis reactions, where the reaction was pseudo-first-order, two separate linear portions occur, one relating to the S_N2 mechanism being dominant, and the second steeper portion relating to the S_N1 mechanism. The L-shaped curvature for solvolysis reactions is due to the S_N1 and S_N2 mechanisms running concurrently with one mechanism being dominant over the other. There were several cases were the experimental S_N2 data departed from linearity, such as *p*-methoxy benzyl chloride. However, the computational calculations performed do not reveal the cause of the departure.

It has been postulated that substrates such as *p*-methoxy benzyl chloride undergo the reaction by a stepwise mechanism in which the rate limiting step is the trapping of a carbocation-leaving group ion-pair by the nucleophile. In order to investigate this further, a quantum mechanical/molecular mechanical model using the GRACE/CHARMM methodology could be used to study the energetics of different rate limiting steps including the trapping of a carbocation-leaving group ion-pair.

13.4.2 Brønsted Correlations

Chrystiuk and Williams concluded that the reaction involving the transfer of the methoxycarbonyl group between isoquinoline and substituted pyridines in aqueous solution was concerted and not stepwise in nature due to the fact that their Brønsted plot followed a 'good' linear relationship over a wide pK_a range. However, the AM1/COSMO theoretical simulation of the system in this thesis concluded that the mechanism was stepwise. A Brønsted plot featuring only the rate determining step gave a reasonably good linear plot which could wrongly be taken as evidence for a concerted mechanism. This shows that great care must be taken when assigning mechanisms on the basis of Brønsted plots. To test the reliability of the COSMO method, a simulation could be used that includes specific solvent-solute interactions by the utilisation of a hybrid QM/MM method. A semi-empirical QM/MM transition-state refinement method such as the GRACE/CHARMM suite of programs could be used to provide an insight into the important specific stabilisation effect of the solvent. The same type of study could be used to investigate the nature of change in mechanisms, if any, using either phosphyl or sulphyl group transfer reactions for which experimental results are available. Different leaving groups or nucleophiles could also be investigated.

13.5 Transition-state models

The calculated results showed that the Pross and Shaik model does not correctly predict all of the effects of changing a wide range of substitutional parameters, whereas the MOFJ model does if great care is taken. The Pross and Shaik model failed due to the apparent incorrect relationship between the barrier height (deformation energy) and transition-state 'looseness'. The calculation of the Pross and Shaik deformation energy neglects the role that the nucleophile plays in a reaction, which must be incorrect.

The participation of solvent stabilisation is unclear in both models. It would be beneficial to examine the extent that the solvent stabilises the reactants, products or transition-state and whether there is preferential stabilisation of one species over the other. The overall effect of solvation can be mimicked by a continuum method such as COSMO. However for the real structural role of solvent stabilisation to become apparent, a QM/MM investigation would have to be carried out.

13.6 Quantum mechanical and molecular mechanical modelling (QM/MM)

Comparison of the semi-empirical and *ab initio* QM/MM models for kinetic isotope effect calculations showed that there is a benefit in using *ab initio* methods. The calculated *ab initio* QM/MM kinetic isotope effects were in excellent agreement with available experimental kinetic isotope effects. However, the use of *ab initio* QM/MM techniques is very computational expensive and therefore its use must be 'chemically justifiable'. Modelling the solvolysis of methoxymethyl chloride suggested that transition-states in QM/MM modelling should be seen as a family of transition-states rather than as a single transition-state structure. This begs the question of how many members are there in a family. The advantage of QM/MM over continuum modelling is that it allows the importance of specific solvation interaction to described when investigating mechanisms. Great care must be taken when initially setting up the system as this can affect the final result. How 'realistic' solvent packing is achieved in the first solvation shell is therefore vitally important and this is an area that would need to be addressed in future work.

13.7 Bader analysis

Baders 'theory of atoms in molecules' showed that cyclopropane, cyclopropanone and oxirane are all closed ring systems in the gas phase, whereas hydroxyoxiranone is an open ring system and should be classed as a zwitterion. Oxiranone was shown to be an intermediate case, using the Bader analysis, as the bonding in the gas phase suggests that it is an α-lactone possessing much ionic character. In solution, cyclopropane, cyclopropanone and oxirane were still all closed ring systems, whereas oxiranone and hydroxyoxiranone are open ring systems, that should be classed as zwitterions and not α-lactones. Hydroxyoxiranone's significantly low strain energy was due to the fact that the open ring structure relieves the ring strain of this system. In general, if the substituent is electron-withdrawing then the zwitterionic is destabilised more than the α-lactone form.

An area of future work involving these ring systems is the inclusion of specific solvent-solute interactions by the utilisation of a hybrid QM/MM method. This would give insight into the specific stabilisation effect of the solvent. As the compounds are

'small', an *ab initio* QM/MM method such as the GRACE/CHARMM/GAMESS-UK suite of programs could be used. The Bader analysis has been used to describe bonding of ground-state species only. Further investigation of reaction mechanisms using Bader analysis could be carried out involving ring compounds or to investigate the degree of ion-pairing that may occur as the leaving group is dissociated in unimolecular nucleophilic substitution reactions.

References

1. a). Foresman, J.B. and Frisch, A. (1996). Exploring Chemistry with Electronic Structure Methods. 2nd Edition. Gaussian, Inc. b). Leach. A.R. (1996). Molecular Modelling - Principles and Applications. Longman.

2. a). Roothaan, C.C.J. (1951). New Developments on Molecular Orbital Theory. *Rev. Mod. Phys.*, **23**, pp. 69-89. b). Hall, G.G. (1951). The Molecular Orbital theory of Chemical Valancy VIII. A Method for Calculating Ionisation Potentials. *Proc. Royal. Soc. London.*, **A205**, pp. 541-552.

3. Møller, C. and Plesset, M.S. (1934). Note on an Approximate Treatment for Many-Electron Systems. *Phys. Rev.*, **46**, pp. 618-622.

4. Pople, J.A., Santry, D.P. and Segal, G.A. (1965). Approximate Self-Consistent Molecular Orbital Theory. I. Invariant Procedures. *J. Chem. Phys.*, **43**, pp. 129-135.

5. Pople, J.A., Beveridge, D.L. and Dobosh, P.A. (1965). Approximate Self-Consistent Molecular Orbital Theory. V. Intermediate Neglect of Differential Overlap. *J. Chem. Phys.*, **47**, pp. 20026-2033.

6. Pople, J.A. and Segal, G.A. (1965). Approximate Self-Consistent Molecular Orbital Theory. II. Calcauations with Complete Negflect of Differential Overlap. *J. Chem. Phys.*, **43**, pp. 136-149.

7. Bingham, R.C., Dewar, M.J.S. and Lo, D.H. (1975). Ground States of Molecules. XXV. MINDO/3. An improved version of the MINDO Semi-empirical SCF-MO Method. *J. Am. Chem. Soc.*, **97**, pp. 1285-1293.

8. Dewar, M.J.S., Zoebisch, E.G., Healy, E.F. and Stewart, J.J.P. (1985). AM1: A New General Purpose Quantum Mechanical Model. *J. Am. Chem. Soc.*, **107**, pp. 3902-3909.

9. Onsager, L. (1936). Electronic Moments of Molecules in Liquid. *J. Am. Chem. Soc.*, **58**, pp. 1486-1493.

10. Miertius, S. and Tomasi, J. (1982). Approximate Evaluations of the Electrostatic Free Energy and Intermediate Energy Changes in Solution Processes. *Chem. Phys.*, **65**, pp. 239-246.

11. Klamt, A. and Schüürmann, G. (1993). COSMO - A New Approach to Dielectric Screening in Solvents with Explicit Expressions for the Screening Energy and its gradient. *J. Chem. Soc. Perkin Trans. 2*, (5), pp. 799-805.

12. Stewart, J.J.P. and Fujitsu Limited. (1993). MOPAC93. Tokyo.

13. Bell, S. and Crighton, J.S. (1984). Locating transition states. *J. Chem. Phys.*, **80**(6), pp. 2464-2475.

14. Flectcher, R. and Reeves, C.M. (1964). *The Computer Journal.* **7**, pp. 149-158.

15. Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S. and Karplus, M. (1983). CHARMM: A program for macromolecular energy, minimisations and dynamics. *J. Comput. Chem.*, **4**(2), pp. 187-217.

16. Cerjan, C.J. and Miller, W.H. (1981). On finding transition states. *J. Chem. Phys.*, **75**(6), pp. 2800-2806.

17. Baker, J. (1986). An Algorithm for the Location of Transition States. *J. Comput. Chem.*, 7, pp. 385-395.

18. Gaussian 94. (1994). Frisch, M.J., Trucks, G.W., Head-Gordon, M., Gill, P.M.W., Wong, M.W., Forman, J.B., Johnson, B.G., Schlegel, H.B., Robb, M.A., Replogle, E.S., Gomperts, R. Andres, J.L., Raghavachari, K. Binkley, J.S., Gonzalez, C., Martin, R.L., Fox, D.J., Defrees, D.J., Baker, J., Stewart, J.J.P. and Pople, J. A. Gaussian, Inc., Pittsburgh, PA.

19. Turner, A.J., Moliner, V. and Williams, I.H. (1999). Transition-State Structural Refinement with GRACE and CHARMM: Flexible QM/MM Modelling for Lactate Dehydrogenase. *Phys. Chem. Chem. Phys.*, 1, pp. 1323-1331.

20. Fischer, S.F. and Karplus, M. (1992). Conjugate peak refinement: an algorithm for finding reaction path and accurate transition states in systems with many degrees of freedom. *Chem. Phys. Lett.*, **194**, pp. 252-261

21. Fukui. K. (1981). The Path of Chemical Reactions - The IRC Approach. *Acc. Chem. Res.*, **14**(12), pp. 363-368

22. Fry. A. (1970). Isotope Effects in Chemical Reactions. ed. Collins, J.C. and Bowman, N.S. Van Nostrand Reinhold company. New York.

23. Shiner, V.J. Jr., Buddenbaum, W.E, Murr, B.L. and Lamaty, G. (1968). Effects of Deuterium Substitution on the Rates of Organic Reactions. XI. α- and β- Deuterium Effects on the Solvolysis Rates of a Series of Substituted 1-Phenylethyl Halides. *J. Am. Chem. Soc.*, **90**(2), pp. 418-427.

24. Williams, I.H. (1984). Theoretical Modelling of Compression Effects in Enzymic Methyl Transfer. *J. Am. Chem. Soc.*, **106**, pp. 7206-7212.

25. Barnes, J.A. and Williams, I.H. (1993). Theoretical Investigation of the Origin of Secondary α-deuterium Kinetic Isotope Effects. *J. Chem. Soc. Chem. Comm.*, **16**, pp 1286-1287.

26. Poirier, R.A., Wang, Y. and Westaway, K.C. (1994). A Theoretical Study of the Relationship between Secondary α-Deuterium Kinetic Isotope Effects and the Structure of S_N2 Transition-States. *J. Am. Chem. Soc.*, **116**, pp. 2526-2533.

27. Glad, S.S. and Jensen, F. (1997). Transition-State Looseness and α-Secondary Kinetic isotope Effects. *J. Am. Chem. Soc.*, **119**, pp. 227-232.

28. a). Williams, I.H. (1983). Force-constant Computations in Cartesian Co-ordinate - Elimination of Translational and Rotational Contributions. *J. Mol. Struct. - THEOCHEM,* **11**(3-4), pp. 275-284. b). Williams, I.H. (1982). On the Representations of Force-fields for Chemically Reacting Systems. *Chem. Phys. Lett.*, **88**(5), pp. 462-466.

29. Page, M. and Williams, A. (1997). Organic and Bio-organic Mechanisms. Addison Wesley Longman.

30. Streitwieser, A.Jr., Hammond, H.A., Jagow, R.H., Williams, R.M., Jesaitis, R.G., Chang, C.J. and Wolf, R. (1970). Acetolysis Reactivities of Substituted Benzyl and Polycyclic Arylmethyl *p*-Toluenesulfonates. Correlations with SCF-π and CNDO MO Methods. *J. Am. Chem. Soc.*, **92**(17), pp. 5141-5150.

31. Castro, E.A., Ibáñez, F., Lagos, S., Schick, M. and Santos, J.G. (1992). Kinetics and Mechanism of the Pyridinolysis of 2,4,6-Trinitrophenyl Acetate and 2,4,6-Trinitrophenyl Methyl Carbonate. *J. Org. Chem.*, **57**, pp. 2691-2693.

32. a) Hughes, E.D. and Ingold, C.K. (1933). Influence of Poles and Polar Linkings on the Course pursued by Elimination Reactions. Part *XV.* Dynamics of the Elimination of Olefins from Quaternary Ammonium Compounds. *J. Chem. Soc.*,

pp. 523-526. b) Hughes, E.D., Ingold, C.K. and Patel, C.S. (1933). Influence of Poles and Polar Linkings on the Course pursued by Elimination Reactions. Part *XVI.* Mechanism of Thermal Decomposition of Quaternary Ammonium Compounds. *J. Chem. Soc.*, pp. 526-531. c) Gleave, J.L., Hughes, E.D. and Ingold, C.K. (1934). Mechanism of Substitution at a Saturated Carbon Atom. Part *III.* Kinetics of the Degradations of Sulphonium Compounds. *J. Chem. Soc.*, pp. 237-244.

33. McMurry, J. (1988). Organic Chemistry. Second Edition. Brooks/Cole Publishing Company. Pacific Grove, California.

34. a) Hughes, E.D. and Ingold, C.K. (1934). Mechanism of Substitution at a Saturated Carbon Atom. Part *IV.* A Discussion of Constitutional and Solvent Effects on the Mechanism, Kinetics, Velocity, and Orientation of Substitution. *J. Chem. Soc.*, pp. 244-255. b) Hughes, E.D. (1934). Mechanism of Substitution at a Saturated Carbon Atom. Part *V.* Hydrolysis of tert.-Butyl Chloride. *J. Chem. Soc.*, pp. 255-258. c) Hughes, E.D., Ingold, C.K. and Shapiro, U.G. (1936). Mechanism of Substitution at a Saturated Carbon Atom. Part *VI.* Hydrolysis of isoPropyl Bromide. *J. Chem. Soc.*, pp. 225-236. d) Cowdrey, W.A., Hughes, E.D., Ingold, C.K., Masterman, S. and Scott. A.D. (1937). Reaction Kinetics and the Walden Inversion. Part VI. Relation of Steric Orientation to Mechanism in Substitution involving Halogen Atoms and Simple or Substituted Hydroxyl Groups. *J. Chem. Soc.*, pp. 1252-1271.

35. a) Dostrovsky, I., Hughes, E.D. and Ingold, C.K. (1945). Mechanism of Substitution at a Saturated Carbon Atom. Part *XXXII.* The Role of Steric Hindrance. (Section G) Magnitude of Steric Effects, Range of Occurrence of Steric and Polar Effects, and Place of the Wagner Rearrangement in Nucleophilic Substitution and Elimination. *J. Chem. Soc.*, pp. 237-244. b) De La Mare, P.B.D., Fowden, L., Hughes, E.D. and Ingold, C.K. (1946). Mechanism of Substitution at a Saturated Carbon Atom. Part *XLIX.* Analysis of Steric and Polar Effects of Alkyl Groups in Bimolecular Nucleophilic Substitution, with Special References to Halogen Exchanges. *J. Chem. Soc.*, pp. 3200-3226.

36. Winstein, S., Grunwald, E. and Jones, H.W. (1951). The Correlation of Solvolysis Rates and the Classification of Solvolysis Reactions into Mechanistic Categories. *J. Am. Chem. Soc.*, **73**, pp. 2700-2707.

37. Bentley, T.W. and Schleyer, P.v.R. (1976). The SN2-SN1 Spectrum. 1. Role of Nucleophilic Solvent Assistance and Nucleophilically Solvated Ion Pair Intermediate in Solvolysis of Primary and Secondary Arenesulfonates. *J. Am. Chem. Soc.*, **98**(24), pp. 7658-7666.

38. Bentley, T.W. and Carter, G.E. (1982). The SN2-SN1 Spectrum. 4. The S_N2 (Intermediate) Mechanism for Solvolyses of *tert*-Butyl Chloride: A Revised Y Scale of Solvent Ionizing Power Based on Solvolyses of 1-Adamantyl Chloride. *J. Am. Chem. Soc.*, **104**, pp. 5741-5747.

39. Bentley, T.W., Bowen, C.T., Morten, D.H. and Schleyer, P.R. (1981). The SN2-SN1 Spectrum. 3. Solvolysis of Secondary and Tertiary Alkyl Sulfonates in Fluorinated Alcohols. Further Evidence for the SN2 (Intermediate) Mechanism. *J. Am. Chem. Soc.*, **103**, pp. 5466-5475.

40. Doering, W.E. and Zeiss, H.H. (1953). Methanolysis of Optically Active Hydrogen 2,4-Dimethylhexyl-4-phthalate. *J. Am. Chem. Soc.*, **75**, pp. 4733-4738.

41. Weiner, H. and Sneen, R.A. (1965). Substitution at a Saturated Carbon Atom. V. A Clarification of the mechanism of Solvolysis of 2-Octyl Sulfonates. Kinetic Considerations. *J. Am. Chem. Soc.*, **87**(2), pp. 292-296.

42. Sneen, R.A. (1973). Ion pairs as Intermediates. Acc. Chem. Res., **6**, 46, pp. 46-53.

43. More O'Ferrall, R.A. (1970). Relationship between *E2* and *E1cB* Mechanisms of β-Elimination. *J. Chem. Soc., B,* pp. 274-277.

44. Jencks, W.P. (1972). General Acid-Base Catalysis of Complex Reactions in Water. *Chem. Rev.*, **72**, pp. 705-719.

45. a) Amyes, T.L. and Richard, J.P. (1990). Concurrent Stepwise and Concerted Substitution Reactions of 4-Methoxybenzyl Derivatives and the Lifetime if the 4-Methoxybenzyl Carbocation. *J. Am. Chem. Soc.*, **112**, pp. 9507-9512. b) Bentley, T.W. and Shim, C.S. (1993). Dual Reaction Channels for Solvolyses of Acyl Chlorides in Alcohol-Water Mixtures. *J. Chem. Soc. Perkin Trans. 2*, pp. 1659-1663. c) Buckley, N. and Oppenheimer, N.J. (1997). Reactions of Charged Substrates. 8. The Nucleophilic Substitution Reactions of (4-Methoxybenzyl)dimethylsulfonium Chloride. *J. Org. Chem.,* 62, pp. 540-551.

46. a) Richard, J.P. and Jencks, W.P. (1984). Concerted Bimolecular Substitution Reactions of 1-Phenylethyl Derivatives. *J. Am. Chem. Soc.*, **106**, pp. 1383-1396. b) Meng, Q. and Thibblin, A. (1995). Solvent-Promoted E2 Reaction Competing with S_N2 Reaction and Stepwise Solvolytic Elimination and Substitution Reactions. *J. Am. Chem. Soc.*, **117**, pp. 9399-9407.

47. Doering, W.E. and Zeiss, H.H. (1953). Methanolysis of Optically Active Hydrogen 2,4-Dimethylhexyl-4-phthalate. *J. Am. Chem. Soc.*, **75**, pp. 4733-4738.

48. Richard, J.P., Amyes, T.L. and Ventour, T. (1991). Absence of Nucleophilic Assistance by Solvent and Azide Ion to the Reaction of Cumyl Derivatives: Mechanism of Nucleophilic Substitution at Tertiary Carbon. *J. Am. Chem. Soc.*, **113**, pp. 5871-5873.

49. Barnes, J.A., Wilkie, J. and Williams, I.H. (1994). Transition-state Structural Variation and Mechanistic Change. *J. Chem. Soc., Faraday Trans.*, **90**(12), pp. 1709-1714.

50. Pauling, L. (1947). Atomic Radii and Interatomic Distances in Metals. *J. Am. Chem. Soc.*, **69**, pp. 542-553.

51. Wilkie, J. and Williams, I.H. (1995). Geometrical preferences for general acid-catalysed hydride transfer study of transition structures for reduction of formaldehyde. *J. Chem. Soc. Perkin Trans. 2*, pp. 1559-1567.

52. Hine, J. (1962). Physical Organic Chemistry, McGraw - Hill Book company, New York.

53. Jencks, W.P. (1980). When Is an Intermediate Not and Intermediate? Enforced Mechanisms of General Acid-Base Catalysed, Carbocation, Carbanion, and Ligand Exchange Reactions. *Acc. Chem. Res.*, **13**, pp. 161-169.

54. Streitwieser, jr. A. (1956). Solvolytic Displacement Reactions at Saturated Carbon Atoms. *Chem. Rev.*, **56**, pp. 571-745.

55. Hine, J., Thomas, C. H. and Ehrenson, S. J. (1955). The Effect of Halogen Atoms on the Reactivity of Other Halogen Atoms in the same Molecule. V. The SN2 Reactivity of Methylene Halides. *J. Am. Chem. Soc.*, **77**, pp. 3886-3889.

56. DeTar, D.L.F., McMullen, D.F. and Luthra, N.P. (1978). Steric Effects in S_N2 Reactions. *J. Am. Chem. Soc.*, **100**, pp. 2484-2493.

57. Ando, T., Tanabe, H. and Yamataka, H. (1984). Kinetic Isotope Effects in the Menschutkin-Type Reaction of Benzyl Benzenesulfonates with *N,N*-Dimethylanilines. Variation in the Transition-State Structure. *J. Am. Chem. Soc.*, **106**, pp. 2084-2092.

58. Knier, B.L. and Jencks, W.P. (1980). Mechanism of Reactions of N-(Methoxymethyl)-N,N-dimethylanilinium Ions with Nucleophilic Reagents. *J. Am. Chem. Soc.*, **102**, pp. 6789-6798.

59. Kazansky, V.B., Senchenya, I.N. and Pankov, A.A. (1991). On the nature of hydrated aliphatic carbenium ions and their role in homogeneous acid catalysis. *J. Mol. Cat.*, **70**, pp. 189-195.

60. Turner, A.J. (1997). *Combined Classical/Quantum Modelling of Chemical Reactions in Enzymes and Solution.* Ph.D. thesis, University of Bath.

61. Tam, J.P., Heath, W.F. and Merrifield, R.B. (1986). Mechanisms for the Removal of Benzyl Protecting Groups in Synthetic Peptides by Trifluoromethanesulfonic Acid-Trifluoroacetic Acid-Dimethyl Sulfide. *J. Am. Chem. Soc.*, **108**, pp. 5242-5251.

62. Maskill, H. and Jencks, W.P. (1987). Solvolysis of Benzyl Azoxytosylate and the Effect of Added Bases and Nucleophiles in Aqueous Trifluoroethanol and Aqueous Aceotonitrile. *J. Am Chem. Soc.*, **109**, pp. 2062-2070.

63. a) Hudson, R.F. and Klopman, G. (1962). Nucleophilic Reactivity. Part II. The Reaction between Substituted Thiophenols and Benzyl Bromides. *J. Chem. Soc.*, pp. 1062-1067. b) Fuchs, R. and Nisbet, A. (1959). Solvent Effects in the Reaction of *p*-Substituted-α-chlorotoluenes with Thiosulfate. The Relationship of Rho and Dielectric Constants. *J. Am Chem. Soc.*, **81**, pp. 2371-2373.

64. a) Grimsrud, E.P. and Taylor, J.W. (1970). Chlorine Kinetic Isotope Effects in Nucleophilic Displacements at a Saturated Carbon. *J. Am Chem. Soc.*, **92**, pp. 739-741. b) Baker, J.W. and Nathan, W.S. (1935). The Mechanism of Aromatic Side-chain Reactions with Special Reference to the Polar Effects of Substitution. Part IV. The Mechanism of Quaternary Salt Formation. *J. Chem. Soc.*, pp. 1840-1844. c) Harris, J.M, Shafer, S.G., Moffatt, J.R. and Becker, A.R. (1979). Predictions of S_N2 Transition State Variation by the Use of More O'Ferrall Plots. *J. Am Chem. Soc.*, **101**, pp. 3295-3300.

65. Willi, A.V., Ho, C.K. and Ghanbarpour, A. (1972). Kinetic α-Deuterium Isotope Effects in the Reactions of Benzyl Chlorides with cyanide Ion and in the Solvolysis of Benzyl Chlorides. *J. Org. Chem.*, **37**(8), pp. 1185-1189.

66. Westaway, K.C. and Waszczylo, Z. (1982). Isotope effects in nucleophilic substitution reactions. IV. The effect of changing a substituent at the α carbon on the structure of the S_N2 transition-states. *Can. J. Chem.*, **60**, pp. 2500-2520.

67. Vitullo, V.P., Grabowski, J. and Sridharan, S. (1980). α-Deuterium Isotope Effects in Benzyl Halides. 2. Reaction of Nucleophiles with Substituted Benzyl Bromides. Evidence for a Change in Transition State Structure with Electron-Donating Substituents. *J. Am. Chem. Soc.*, **102**, pp. 6463-6465.

68. Östman, B. (1965). Secondary α-Deuterium Isotope Effects and Relative Rates in the Halogen Exchange Reaction of Benzyl and Thenyl Chlorides. *J. Am. Chem. Soc.*, **87**, 14, pp. 3163-3167.

69. Hammett, L.P. (1935). Some Relations Between Reaction Rates and Equilibrium Constants. *Chem. Rev.*, *17*, pp. 125-136.

70. Young, P.R. and Jencks, W.P. (1979). Separation of Polar and Resonance Substituent Effects in the Reactions of Acetophenones with Bisulfite and Benzyl Halides with Nucleophiles. *J. Am. Chem. Soc.*, **101**(12), pp. 3288-3294.

71. Hammond, G. S., Reeder, C.E., Fang, F.T. and Kochi, J.K. (1958). Mountaineering Across the Andes. *J. Am. Chem. Soc.*, **80**, pp. 568-580.

72. Bordwell, F.G. and Hughes, D.L. (1986). Rate-Equilibrium Relationships for Reactions of Families of Carbanion Nucleophiles with *N*-Benzyl-*N*-*N*-dimethylanilinium Cations and with Alkyl Chloride, Bromides, and Iodides. *J. Am. Chem. Soc.*, **108**, pp. 7300-7309.

73. Ballisteri, F.P., Maccarone, E. and Mamo, A. (1976). Kinetics and Mechanism of Benzylation of Anilines. *J. Org. Chem.*, **41**, pp. 3364-3367.

74. Yukawa, Y. and Tsuno. Y. (1959). Mutlicomponent Rate-Equilibrium Relationships. *Bull. Chem Soc. Jpn.*, **32**, pp. 965-971.

75. Chapman, N.B. and Shorter, J. (1978). Correlation Analysis in Chemistry, Plenum Press, New York.

76. Unpublished results from J. P. Richard, University at Buffalo, State University of New York. [personal communication to Williams, I.H.]

77. Williams, A. (1994). The Diagnosis of Concerted Organic Mechanisms. *Chem. Soc. Rev.*, **23**, pp. 93-100.

78. Chrystiuk, E. and Williams, A. (1987). A Single transition state in the Transfer of the Methoxycarbonyl Group between Isoquinoline and substituted Pyridine in Aqueous Solution. *J. Am. Chem. Soc.*, **109**, pp. 3040-3046.

79. McClelland, R.A. and Santry, L.J. (1983). Reactivity of Tetrahedral Intermediates. *Acc. Chem. Res.*, **16**, pp. 394-399.

80. Williams, A. (1989). Concerted Mechanisms of Acyl Group Transfer Reactions in Solution. *Acc. Chem. Res.*, **22**, pp. 387-392.

81. Bender, M.L. (1960). Mechanisms of Catalysis of Nucleophilic Reactions of Carboxylic Acid Derivatives. *Chem. Rev.*, **60**, pp. 53-113.

82. Jencks, W.P. (1981). Ingold Lecture: How Does a Reaction Choose Its Mechanism? *Chem. Soc. Rev.*, **10**, pp. 345-375.

83. Williams, I.H. (1988). Theoretical probes for activated-complex structures and properties: specific solvation effects and a simulated Brønsted correlation. *Bull. Soc. Chim. Fr.* pp. 192-198.

84. Hammond, R.B. and Williams, I.H. (1989). Brönsted Exponents and Activated-complex Structure: An AM1 SCF-MO Theoretical Simulation of Rate-Equilibrium Correlation for the Transfer of the Methoxycarbonyl Group between Isoquinoline and Substituted Pyridine. *J. Chem. Soc. Perkin Trans. 2*, pp. 59-66.

85. Bourne, N. and Williams, A. (1984). Evidence for a Single Transition State in the Transfer of the Phosphoryl Group ($-PO_3^{2-}$) to Nitrogen Nucleophiles from Pyridino-*N*-phosphonates. *J. Am. Chem. Soc.*, **106**, pp. 7591-7596.

86. Skoog, M.T. and Jencks, W.P. (1984). Reactions of Pyridines and Amines N-Phosphorylated Pyridines. *J. Am. Chem. Soc.*, **106**, pp. 7597-7606.

87. Williams, A. (1985). Bonding in Phosphoryl ($-PO_3^{2-}$) and Sulfuryl ($-SO_3^-$) Group Transfer between Nitrogen Nucleophiles as Determined from Rate constants for Identity Reactions. *J. Am. Chem. Soc.*, **107**, pp. 6335-6339.

88. Davidson, W.R., Sunner, J. and Kebarle, P. (1979). Hydrogen Bonding of Water to Onium Ions. Hydration of Substituted Pyridinium Ions and Related Systems. *J. Am. Chem. Soc.*, **101**(7), pp. 1675-1680.

89. Arnett, E. M., Chawla, B., Bell, Taagepera, L., Hehre, W.J. and Taft, R.W. (1977). Solvation and Hydrogen Bonding of Pyridinium Ions. *J. Am. Chem. Soc.*, **99**(17), pp. 5729-5738.

90. Pross, A. and Shaik, S.S. (1981). Reactivity-Selectivity Relationships. A Quantum Mechanical Approach to Transition State Structure. Application to S_N2 Reaction of Benzyl Derivatives. *J. Chem. Soc.*, **103**, pp. 3702-3709.

91. Lowry, T.H. and Richardson, K.S. (1987). Mechanism and Theory on organic Chemistry. 3rd Edition, Harper and Row, Publishers, New York.

92. Pross, A. (1995). Theoretical & Physical Principles of Organic Reactivity. Wiley-Interscience Publications.

93. Shaik, S.S., Schlegal, H.B. and Wolfe, S. (1992). Theoretical aspects of physical organic chemistry: the S_N2 mechanism. Wiley-Interscience Publications.

94. Shaik, S.S. (1985). The Collage of S_N2 Reactivity Patterns: A State Correlation Diagram Model. *Prog. Phys. Org. Chem.*, **15**, pp. 197-337.

95. Mitchell, D.J., Schlegal, H.B., Shaik, S.S. and Wolfe, S. (1985). Relationships between geometries and energies of identity S_N2 transition states: the dominant role of the distortion energies and its origin. *Can. J. Chem.*, **63**, pp. 1642-1648.

96. Shaik, S.S. (1988). How is Transition State Looseness Related to the Reaction Barrier? *J. Am. Chem. Soc.*, **110**, pp. 1127-1131.

97. Pross, A. and Shaik, S.S. (1982). S_N2 Reactivity of CH_3X Derivatives. A Valence Bond Approach. *J. Chem. Soc.*, **104**, pp. 2708-2719.

98. Pross, A. and Shaik, S.S. (1983). A Qualitative Valence Bond Approach to Organic Reactivity. *Acc. Chem. Res.*, **16**, pp. 363-370.

99. Barnes, J.A. and Williams, I.H. (1996). Theoretical modelling of kinetic isotope effects for glycoside hydrolysis in aqueous solution by a hybrid quantum-mechanical /molecular mechanical method. *Chem. Commun.*, 2, pp. 193-194.

100. Jorgensen, W.L., Chandrasekhar, J., Madura, J.D, Impey, R.W. and Klein, M.L. (1983). Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.*, **79**(2), pp. 926-935.

101. Ousterhout, J.K. (1994). Tcl and Tk toolkit. Addison Wesley ISBN 0-201-6337-X.

102. XMol version 1.3.1. dba Minnesota Supercomputer Center Inc.

103. Sayle, R. RasMol v2.5: A molecular visualisation Program. Biomolecular Structure, Glaxo Research and Development, Greenford, Middlesex, UK.

104. GAMESS-UK. (1995-7). Daresbury Laboratories. Guest, M. F., Van Lenthe, J. H., Schoffel, K., Sherwood, P. and Harrison, R. J. Contributors: Amos, R. D., Buenker, R. J., Dupuis, M., Handy, N. C., Hillier, I. H., Knowles, P. J., Bonacic-Koutecky, V., Von Niessen, W., Saunders, V. R. and Stone, A. J. The package is derived from the original GAMESS code due to Dupuis, M., Spangler, D. and Wendoloski, J.

105. Field, M.J. Bash, P.A. and Karplus, M. (1990). A Combined Quantum Mechanical and Molecular Mechanical Potential for Molecular Dynamics Simulation. *J. Comp. Chem.,* 11(6), pp. 700-733.

106. Hartsough, D.S. and Merz, Jr., K.M. (1995). Potential of Mean Force Calculations on the S_N1 Fragmentation of tert-Butyl Chloride. *J. Phys Chem.*, **99**, 384-390.

107. Toteva, M.M. and Richard, J.P. (1996). Mechanism for Nucleophilic Substitution and elimination Reactions at Tertiary Carbon in Largely Aqueous Solutions: Lifetime of a Simple Tertiary Carbocation. *J. Am. Chem. Soc.*, **118**, pp. 11434-11445.

108. Powell, M.J.D. (1971). Math. Programming, 1, 26.

109. a) Llewellyn, J.A., Robertson, R.E. and Scott. J.M. (1960). Some deuterium isotope effects: I. Water solvolysis of methyl-d, esters. *Can. J. Chem.*, **38**, pp. 222-232. b) Leffek, K.T., Llewellyn, J.A. and Robertson, R.E. (1960). Some deuterium isotope effects: II. α-Deuterium effects in the water solvolysis of some alkyl compounds. *Can. J. Chem.*, **38**, pp. 1504-1510. c) Leffek, K.T., Llewellyn, J.A. and Robertson, R.E. (1960). Some deuterium isotope effects: IV. β-

Deuterium effects in the water solvolysis of ethyl, isopropyl, and tert-butyl compounds. *Can. J. Chem.*, **38**, pp. 2171-2177.

110. Craze, G. and Kirby, A.J. (1978). Bimolecular Nucleophilic Substitution on an Acetal. *J. Chem. Soc. Perkin Trans. 2,* pp. 357-368.

111. Fry. A. (1970). Isotope Effects in Chemical Reactions. ed. Collins, J.C. and Bowman, N.S. Van Nostrand Reinhold company. New York.

112. Winstein, S., Grunwald, E. and Jones, H.W. (1951). The Correlation of Solvolysis Rates and the Classification of Solvolysis Reactions into Mechanistic Categories. *J. Am. Chem. Soc.*, **73**, pp. 2700-2707.

113. Jones, T.C. and Thornton, E.R. (1967). Solvolysis Mechanisms. S_N1-Like Behaviour of Methyl Chloromethyl Ether. Sensitivity to Solvent Ionising Power and α-Deuterium Isotope Effect. *J. Am. Chem. Soc.*, **89**, pp. 4863-4867.

114. Ballinger, P., De La Mare, P.D.B., Kohnstam, G. and Prestt, B.M. (1955). The Reaction of Chlorodimethyl Ether with Ethanol and with Ethoxide Ions. *J. Am. Chem. Soc.*, pp. 3641-3647.

115. Craze, G. and Kirby, A.J. (1978). Bimolecular Nucleophilic Substitution on an Acetal. *J. Chem. Soc. Perkin Trans. 2,* pp. 357-368.

116. Knier, B.L. and Jencks, W.P. (1980). Mechanism of Reactions of N-(Methoxymethyl)-N,N-dimethylanilinium Ions with Nucleophilic Reagents. *J. Am. Chem. Soc.*, **102**, pp. 6789-6798.

117. Sinnott, M.L. (1984) The Chemistry of Enzyme Action, Ed. Page, M.I. Elsevier.

118. Garrett, E.R. and Mehta, P.J. (1972). Solvolysis of Adenine Nucleosides. I. Effects of Sugars and Adenine Substituents on Acid Solvolyses. *J. Am. Chem. Soc.*, **94**, pp. 8532-8541.

119. Parkin, D.W. and Schramm, V.L. (1987). Catalytic and Allosteric Mechanism of AMP Nucleosidase from Primary, β-Secondary, and Multiple Heavy Atom Kinetic Isotope Effects. *Biochem.*, **26**, pp. 913-920.

120. Jones, T.C. and Thornton, E.R. (1967). Solvolysis Mechanisms. S_N1-Like Behaviour of Methyl Chloromethyl Ether. Sensitivity to Solvent Ionising Power and α-Deuterium Isotope Effect. *J. Am. Chem. Soc.*, **89**, pp. 4863-4867.

121. Westheimer, F.H. (1987). Why Nature Chose Phosphates. *Science.* 235, pp. 1173-1178.

122. L'abbé, G. (1980). Heterocyclic Analogues of Methylenecyclopropanes. *Angew. Chem. Int. Ed. Engl.*, **19**, pp. 276-289.

123. Cowdrey, W.A., Hughes, E.D., Ingold, C.K., Masterman, S. and Scott, A.D. (1937). Reaction Kinetics and the Walden Inversion. Part VI. Relation of Steric Orientation to Mechanism in Substitution involving Halogen Atoms and Simple or Substituted Hydroxyl Groups. *J. Chem. Soc.,* pp. 1252-1271.

124. Hammett, L.P. *Physical Organic Chemistry*, McGraw-Hill, New York, 1940.

125. Grunwald, E. and Winstein, S. (1948). The role of Neighbouring Groups in replacement Reactions. XV. Rates and Medium Effects in the Alcoholysis and Hydrolysis of α-Bromopropionate Ion. The Carboxylate Ion group. *J. Am. Chem. Soc.*,**70**, pp. 841-846.

126. Rodriquez, C.F. and Williams I.H. (1997). Ring-strain and enthalpy of formation of oxiranone: an *ab initio* theoretical determination. *J. Chem. Soc., Perkin Trans 2*, **5**, pp. 953-957.

127. Firth-Clark, S., Rodriquez, C.F. and Williams I.H. (1997). Hydroxyoxiranone: an *ab initio* MO investigation of the structure and stability of a model for a possible α-lactone intermediate in hydrolysis of sialyl glycosides. *J. Chem. Soc., Perkin Trans 2*, **10**, pp. 1943-1948.

128. Rodriquez, C. F. and Williams I. H. (1997). *Ab initio* theoretical investigation of the mechanism for α-lactone formation from α-halocarboxylates: leaving group, substituent, solvent and isotope effect. *J. Chem. Soc., Perkin Trans 2*, **5**, pp. 959-965.

129. a) Bader, R.F.W. (1990). *Atoms in Molecules: A Quantum Theory*, Oxford University Press: Oxford, UK. b) Bader, R.F.W., Popelier, P.L.A. and Keith, T.A. (1994). Theoretical Definition of a Functional Group and the Molecular Orbital Paradigm. *Angew. Chem. Int. Ed. Engl.*, **33**, pp. 620-631.

130. Krafa, E. and Cremer, D. (1985). Theoretical Determination of Molecular Structure and conformation. 15. Three-membered rings: bent bonds, ring-strain, and surface delocalisation. *J. Am. Chem. Soc.*, **107**, pp. 3800-3810.

131. Krafa, E. and Cremer, D. (1985). Theoretical Determination of Molecular Structure and conformation. 16. Substituted cyclopropanes - an electron density model of substituent ring interactions. *J. Am. Chem. Soc.,* **107**, pp. 3810-3817.

132. Howard, S.T., Foreman, J.P. and Edwards, P.G. (1996). Electronic Structure of Aryl- and Alkylphosphines. *Inorg. Chem.,* **35**, pp. 5805-5812.

133. Cremer, D. and Kraka, E. (1988). *Structure and Reactivity*, eds. J. F. Liebman and A. Greenberg, VCH, New York.

134. Krafa, E. and Cremer, D. (1992). Description of chemical reactions in terms of the properties of the electron density. *J. Mol. Struct. (THEOCHEM)*, **255**, pp. 189-206.

135. Popelier, P.L.A. (1998). Characterisation of a Dihydrogen bond on the basis of the electron density. *J. Phys. Chem. A*, **102**, pp. 1873-1878.

Appendix A

Data for Chapter 4 Energy Surfaces

Appendix A - Data for Chapter 4 Energy Surfaces

The data for the energy surfaces in Chapter 4 are given below for, $RCl + H_2O \rightarrow ROH_2^+ + Cl^-$, where R= methyl, ethyl, ^{iso}propyl and ^{t}butyl. The first and second columns are the bond orders and the third column is the energy in kJ mol^{-1}.

Gas-phase energy surfaces

Methyl

1.0	1.0	568.7	0.5	0.6	343.8	1.0	1.0	589.4	0.5	0.6	305.5
0.9	1.0	527.2	0.4	0.6	334.1	0.9	1.0	543.9	0.4	0.6	284.5
0.8	1.0	491.6	0.3	0.6	329.1	0.8	1.0	503.8	0.3	0.6	269.1
0.7	1.0	450.1	0.2	0.6	335.1	0.7	1.0	456.2	0.2	0.6	264.4
0.6	1.0	404.4	0.1	0.6	345.9	0.6	1.0	402.8	0.1	0.6	267.2
0.5	1.0	361.9	0.0	0.6	363.7	0.5	1.0	352.2	0.0	0.6	286.2
0.4	1.0	321.0	0.5	0.5	343.4	0.4	1.0	301.9	0.5	0.5	295.2
0.3	1.0	295.6	0.4	0.5	343.8	0.3	1.0	267.0	0.4	0.5	283.2
0.2	1.0	290.7	0.3	0.5	346.0	0.2	1.0	251.6	0.3	0.5	274.2
0.1	1.0	298.4	0.2	0.5	355.8	0.1	1.0	248.2	0.2	0.5	273.4
0.0	1.0	313.8	0.1	0.5	367.8	0.0	1.0	262.3	0.1	0.5	277.9
0.9	0.9	490.6	0.0	0.5	387.0	0.9	0.9	502.7	0.0	0.5	298.7
0.8	0.9	460.0	0.4	0.4	354.9	0.8	0.9	467.3	0.4	0.4	281.0
0.7	0.9	425.0	0.3	0.4	365.9	0.7	0.9	425.9	0.3	0.4	279.1
0.6	0.9	386.7	0.2	0.4	380.1	0.6	0.9	379.6	0.2	0.4	282.8
0.5	0.9	351.1	0.1	0.4	393.9	0.5	0.9	335.6	0.1	0.4	289.2
0.4	0.9	316.3	0.0	0.4	414.9	0.4	0.9	291.2	0.0	0.4	311.3
0.3	0.9	294.9	0.3	0.3	385.0	0.3	0.9	260.5	0.3	0.3	283.3
0.2	0.9	292.4	0.2	0.3	403.6	0.2	0.9	247.1	0.2	0.3	290.7
0.1	0.9	300.6	0.1	0.3	419.3	0.1	0.9	244.9	0.1	0.3	297.4
0.0	0.9	316.2	0.0	0.3	442.6	0.0	0.9	260.0	0.0	0.3	322.2
0.8	0.8	434.5	0.2	0.2	425.2	0.8	0.8	436.8	0.2	0.2	301.2
0.7	0.8	406.2	0.1	0.2	442.3	0.7	0.8	401.7	0.1	0.2	307.3
0.6	0.8	375.3	0.0	0.2	467.5	0.6	0.8	362.6	0.0	0.2	326.1
0.5	0.8	346.6	0.1	0.1	457.1	0.5	0.8	325.3	0.1	0.1	317.3
0.4	0.8	318.2	0.0	0.1	483.5	0.4	0.8	287.0	0.0	0.1	334.1
0.3	0.8	301.0	0.0	0.0	495.0	0.3	0.8	260.5	0.0	0.0	356.0
0.2	0.8	300.7				0.2	0.8	249.1			
0.1	0.8	309.6				0.1	0.8	248.2			
0.0	0.8	325.5				0.0	0.8	264.3			
0.7	0.7	386.0				0.7	0.7	374.5			
0.6	0.7	364.6				0.6	0.7	344.5			
0.5	0.7	344.6				0.5	0.7	315.6			
0.4	0.7	324.8				0.4	0.7	285.5			
0.3	0.7	312.9				0.3	0.7	263.9			
0.2	0.7	315.6				0.2	0.7	255.5			
0.1	0.7	325.3				0.1	0.7	256.3			
0.0	0.7	341.9				0.0	0.7	273.7			
0.6	0.6	353.8				0.6	0.6	324.8			

Ethyl

^{iso}Propyl						^{t}Butyl					
1.0	1.0	599.4	0.4	0.6	247.8	1.0	1.0	650.9	0.4	0.6	234.3
0.9	1.0	552.0	0.3	0.6	224.9	0.9	1.0	599.9	0.3	0.6	200.7
0.8	1.0	510.5	0.2	0.6	215.8	0.8	1.0	553.1	0.2	0.6	177.5
0.7	1.0	461.5	0.1	0.6	218.5	0.7	1.0	496.0	0.1	0.6	168.8
0.6	1.0	404.5	0.0	0.6	232.0	0.6	1.0	429.8	0.0	0.6	177.1
0.5	1.0	353.7	0.5	0.5	259.3	0.5	1.0	365.3	0.5	0.5	249.4
0.4	1.0	301.2	0.4	0.5	239.4	0.4	1.0	297.3	0.4	0.5	219.6
0.3	1.0	261.9	0.3	0.5	221.4	0.3	1.0	243.7	0.3	0.5	193.8
0.2	1.0	241.2	0.2	0.5	215.0	0.2	1.0	207.4	0.2	0.5	172.7
0.1	1.0	241.2	0.1	0.5	218.5	0.1	1.0	183.8	0.1	0.5	168.1
0.0	1.0	251.3	0.0	0.5	233.3	0.0	1.0	173.6	0.0	0.5	186.2
0.9	0.9	507.7	0.4	0.4	226.5	0.9	0.9	553.2	0.4	0.4	198.5
0.8	0.9	469.5	0.3	0.4	217.0	0.8	0.9	511.4	0.3	0.4	176.8
0.7	0.9	425.3	0.2	0.4	213.9	0.7	0.9	460.9	0.2	0.4	165.7
0.6	0.9	376.3	0.1	0.4	218.3	0.6	0.9	402.3	0.1	0.4	165.9
0.5	0.9	329.3	0.0	0.4	233.0	0.5	0.9	343.2	0.0	0.4	166.1
0.4	0.9	281.4	0.3	0.3	211.3	0.4	0.9	280.1	0.3	0.3	163.9
0.3	0.9	245.4	0.2	0.3	211.3	0.3	0.9	229.4	0.2	0.3	156.8
0.2	0.9	227.6	0.1	0.3	216.5	0.2	0.9	194.6	0.1	0.3	158.9
0.1	0.9	228.4	0.0	0.3	210.7	0.1	0.9	173.3	0.0	0.3	175.2
0.0	0.9	238.9	0.2	0.2	212.2	0.0	0.9	170.5	0.2	0.2	153.9
0.8	0.8	435.1	0.1	0.2	217.1	0.8	0.8	472.2	0.1	0.2	155.3
0.7	0.8	395.8	0.0	0.2	230.0	0.7	0.8	425.5	0.0	0.2	170.6
0.6	0.8	352.8	0.1	0.1	221.5	0.6	0.8	371.9	0.1	0.1	153.6
0.5	0.8	311.8	0.0	0.1	237.1	0.5	0.8	319.3	0.0	0.1	166.6
0.4	0.8	268.7	0.0	0.0	253.4	0.4	0.8	263.0	0.0	0.0	178.9
0.3	0.8	236.2				0.3	0.8	218.8			
0.2	0.8	221.1				0.2	0.8	187.3			
0.1	0.8	222.5				0.1	0.8	169.6			
0.0	0.8	233.6				0.0	0.8	163.6			
0.7	0.7	351.2				0.7	0.7	387.6			
0.6	0.7	301.2				0.6	0.7	342.6			
0.5	0.7	275.8				0.5	0.7	298.1			
0.4	0.7	247.8				0.4	0.7	249.7			
0.3	0.7	224.9				0.3	0.7	211.0			
0.2	0.7	215.8				0.2	0.7	182.0			
0.1	0.7	218.5				0.1	0.7	168.7			
0.0	0.7	232.0				0.0	0.7	174.5			
0.6	0.6	301.2				0.6	0.6	309.2			
0.5	0.6	275.8				0.5	0.6	275.1			

Aqueous-phase energy surfaces

Methyl

1.0	1.0	268.3	0.7	0.2	-3.5
1.0	0.9	225.8	0.7	0.1	-0.3
1.0	0.8	188.5	0.7	0.0	-10.0
1.0	0.7	144.0	0.6	0.6	52.1
1.0	0.6	94.0	0.6	0.5	40.8
1.0	0.5	46.6	0.6	0.4	28.0
1.0	0.4	-0.9	0.6	0.3	20.1
1.0	0.3	-32.2	0.6	0.2	22.4
1.0	0.2	-44.2	0.6	0.1	27.2
1.0	0.1	-43.4	0.6	0.0	17.0
1.0	0.0	-50.1	0.5	0.5	41.1
0.9	0.9	189.8	0.5	0.4	39.9
0.9	0.8	158.3	0.5	0.3	41.0
0.9	0.7	121.5	0.5	0.2	48.6
0.9	0.6	79.3	0.5	0.1	54.7
0.9	0.5	38.8	0.5	0.0	41.2
0.9	0.4	-1.7	0.4	0.4	51.2
0.9	0.3	-28.6	0.4	0.3	62.5
0.9	0.2	-37.5	0.4	0.2	75.7
0.9	0.1	-35.2	0.4	0.1	83.6
0.9	0.0	-43.1	0.4	0.0	65.7
0.8	0.8	133.5	0.3	0.3	81.5
0.8	0.7	104.1	0.3	0.2	100.5
0.8	0.6	70.5	0.3	0.1	109.3
0.8	0.5	37.5	0.3	0.0	87.1
0.8	0.4	3.9	0.2	0.2	122.0
0.8	0.3	-18.5	0.2	0.1	108.2
0.8	0.2	-24.7	0.2	0.0	98.7
0.8	0.1	-22.1	0.1	0.1	141.0
0.8	0.0	-30.1	0.1	0.0	114.4
0.7	0.7	84.4	0.0	0.0	89.3
0.7	0.6	62.2			
0.7	0.5	39.0			
0.7	0.4	14.9			
0.7	0.3	-0.9			

Ethyl

1.0	1.0	302.2	0.2	0.7	-35.9
0.9	1.0	255.8	0.1	0.7	-37.0
0.8	1.0	215.0	0.0	0.7	-45.5
0.7	1.0	165.2	0.6	0.6	41.8
0.6	1.0	108.3	0.5	0.6	23.0
0.5	1.0	53.6	0.4	0.6	1.8
0.4	1.0	-1.8	0.3	0.6	-13.9
0.3	1.0	-40.8	0.2	0.6	-19.3
0.2	1.0	-60.7	0.1	0.6	-18.9
0.1	1.0	-67.6	0.0	0.6	-27.4
0.0	1.0	-75.2	0.5	0.5	15.2
0.9	0.9	215.8	0.4	0.5	4.6
0.8	0.9	180.4	0.3	0.5	-3.0
0.7	0.9	138.1	0.2	0.5	-3.5
0.6	0.9	89.1	0.1	0.5	-1.6
0.5	0.9	41.8	0.0	0.5	-11.4
0.4	0.9	-7.7	0.4	0.4	5.6
0.3	0.9	-42.4	0.3	0.4	6.2
0.2	0.9	-59.5	0.2	0.4	11.6
0.1	0.9	-63.0	0.1	0.4	14.7
0.0	0.9	-70.6	0.0	0.4	2.1
0.8	0.8	151.1	0.3	0.3	13.5
0.7	0.8	115.6	0.2	0.3	23.8
0.6	0.8	75.7	0.1	0.3	27.9
0.5	0.8	34.3	0.0	0.3	13.9
0.4	0.8	-6.9	0.2	0.2	35.4
0.3	0.8	-37.0	0.1	0.2	40.4
0.2	0.8	-50.5	0.0	0.2	25.1
0.1	0.8	-54.2	0.1	0.1	47.9
0.0	0.8	-63.1	0.0	0.1	31.0
0.7	0.7	89.7	0.0	0.0	14.5
0.6	0.7	59.9			
0.5	0.7	29.5			
0.4	0.7	-3.6			
0.3	0.7	-26.8			

^{iso}Propyl						^{t}Butyl					
1.0	1.0	329.0	0.2	0.6	-45.4	1.0	1.0	-255.1	0.2	0.6	-242.1
0.9	1.0	283.0	0.1	0.6	-44.7	0.9	1.0	-254.2	0.1	0.6	-247.8
0.8	1.0	240.5	0.0	0.6	-52.75	0.8	1.0	-253.3	0.0	0.6	-257.4
0.7	1.0	190.9	0.5	0.5	-1.10	0.7	1.0	-257.7	0.5	0.5	-240.9
0.6	1.0	134.6	0.4	0.5	-19.38	0.6	1.0	-259.6	0.4	0.5	-240.5
0.5	1.0	80.4	0.3	0.5	-33.44	0.5	1.0	-261.9	0.3	0.5	-241.4
0.4	1.0	23.6	0.2	0.5	-40.4	0.4	1.0	-265.2	0.2	0.5	-240.2
0.3	1.0	-19.6	0.1	0.5	-38.6	0.3	1.0	-268.5	0.1	0.5	-242.3
0.2	1.0	-44.1	0.0	0.5	-44.45	0.2	1.0	-267.6	0.0	0.5	-258.7
0.1	1.0	-51.7	0.4	0.4	-26.8	0.1	1.0	-269.8	0.4	0.4	-237.9
0.0	1.0	-59.1	0.3	0.4	-34.93	0.0	1.0	-268.9	0.3	0.4	-234.7
0.9	0.9	242.7	0.2	0.4	-37.01	0.9	0.9	-251.8	0.2	0.4	-235.9
0.8	0.9	200.1	0.1	0.4	-33.08	0.8	0.9	-252.8	0.1	0.4	-239.2
0.7	0.9	155.9	0.0	0.4	-43.01	0.7	0.9	-258.2	0.0	0.4	-240.5
0.6	0.9	106.0	0.3	0.3	-35.02	0.6	0.9	-261.6	0.3	0.3	-231.5
0.5	0.9	57.8	0.2	0.3	-33.9	0.5	0.9	-263.0	0.2	0.3	-231.8
0.4	0.9	7.131	0.1	0.3	-37.43	0.4	0.9	-264.1	0.1	0.3	-233.6
0.3	0.9	-31.5	0.0	0.3	-43.14	0.3	0.9	-265.2	0.0	0.3	-253.4
0.2	0.9	-52.8	0.2	0.2	-30.82	0.2	0.9	-264.9	0.2	0.2	-228.6
0.1	0.9	-57.7	0.1	0.2	-34.19	0.1	0.9	-262.4	0.1	0.2	-228.8
0.0	0.9	-65.6	0.0	0.2	-39.38	0.0	0.9	-267.2	0.0	0.2	-251.7
0.8	0.8	168.5	0.1	0.1	-31.4	0.8	0.8	-245.0	0.1	0.1	-228.6
0.7	0.8	127.8	0.0	0.1	-36.82	0.7	0.8	-253.0	0.0	0.1	-244.7
0.6	0.8	84.8	0.0	0.0	-47.6	0.6	0.8	-253.8	0.0	0.0	-268.4
0.5	0.8	43.3				0.5	0.8	-255.7			
0.4	0.8	-1.3				0.4	0.8	-257.4			
0.3	0.8	-35.3				0.3	0.8	-259.0			
0.2	0.8	-54.35				0.2	0.8	-257.8			
0.1	0.8	-57.91				0.1	0.8	-254.7			
0.0	0.8	-64.2				0.0	0.8	-255.9			
0.7	0.7	58.209				0.7	0.7	-251.1			
0.6	0.7	33.756				0.6	0.7	-251.2			
0.5	0.7	-52.21				0.5	0.7	-251.5			
0.4	0.7	-45.63				0.4	0.7	-252.4			
0.3	0.7	-55.4				0.3	0.7	-252.5			
0.2	0.7	-54.7				0.2	0.7	-250.3			
0.1	0.7	-52.75				0.1	0.7	-249.1			
0	0.7	-62.75				0.0	0.7	-258.4			
0.6	0.6	38.209				0.6	0.6	-241.5			
0.5	0.6	13.756				0.5	0.6	-246.4			
0.4	0.6	-12.21				0.4	0.6	-244.4			
0.3	0.6	-35.63				0.3	0.6	-243.			

Appendix B

GRACE

Appendix B- GRACE

Overview

GRACE can be thought of as an acronym for General Reprogrammable Algebraic Chemistry Engine. The GRACE software package was developed by a Ph.D. student of the Williams' research group, A. J. Turner. It provides a wide range of functionality for the location and characterisation of transition structures in hybrid QM/MM treatments of large, flexible systems involving several thousand degrees of freedom. It may be used in conjunction with CHARMM to perform semi-empirical QM/MM calculations or as an interface between CHARMM and CADPAC or GAMESS-UK for *ab initio* QM/MM calculations.

GRACE employs the 'eigenvector-following' method for the location of transition structures in association with a subset-Hessian approach (see Chapter 9). It permits calculation of the vibrational frequencies for the 'core' within the 'environment' of a larger system. The distinction between core and environment in this sense is completely separate from the choice of the QM and MM subsystems. GRACE can also perform intrinsic reaction co-ordinate calculations from a first-order saddle-point, and thus allows rigorous characterisation of a transition-state for very large systems.

GRACE contains a 'flexible free format file parser extract' for importation of data from other programs. In conjunction with the native file and list manipulation tools in TCL, this allows it to overcome many of the translation and compatibility problems faced by modern computational chemists. As well as having a z-matrix to Cartesian co-ordinate converter, GRACE can autogenerate valence co-ordinates, and translate these into CHARMM co-ordinates, for analysis and constraints use. This system has also interfaced with I. H. Williams' CAMVIB and CAMISO vibrational analysis programs. GRACE contains routines to fit MM point charges to the QM/MM coloumbic term and resolve lone-pairs which allows more realistic interpolation of QM/MM functions by pure MM functions.

The current features in GRACE:

- Semi-empirical QM/MM
- *Ab initio* QM/MM
- Transition-state searching
- Intrinsic reaction co-ordinate
- Internal co-ordinate generation
- Lone pair location
- Hessian computation
- Charge fitting
- Frequency calculations
- Kinetic isotope effect calculations
- Structured programming language for future developmental additions

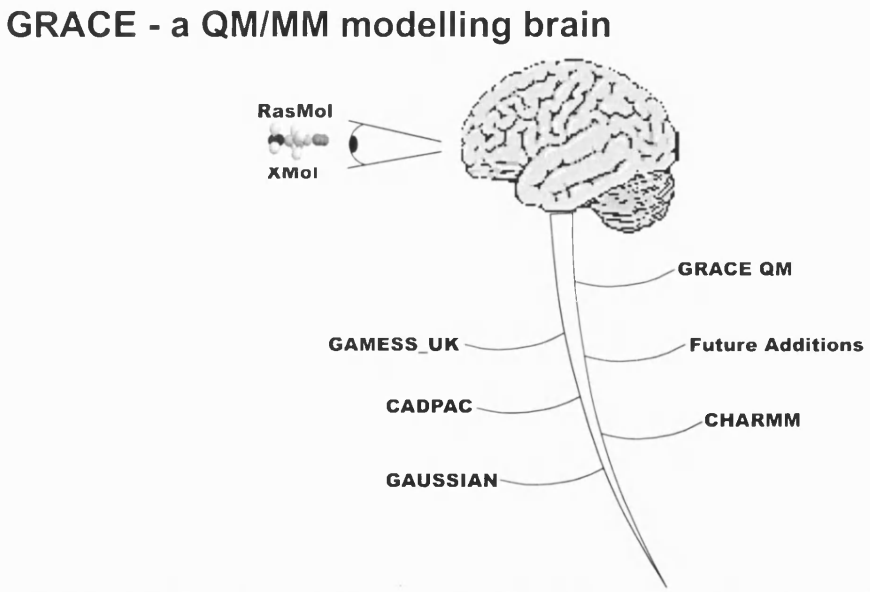


Figure B.1: GRACE is based on a mathematical program whose functions have been oriented to the needs of the computational chemist.

GRACE can be viewed as a QM/MM modelling brain, Figure B.1. The code contains very few internal energy functions, with the majority functions coming from external codes. The QM/MM and visualisation codes are not an in-built component of GRACE, however the algorithms and structures in GRACE allow control over external programs to facilitate a synergy between them, and allow novel modelling approaches. GRACE-QM is a possible future extension of GRACE which could have a native semi-empirical QM/MM method instead of using that from CHARMM.

TCL - the language of GRACE

The main interface of GRACE is via a tool control language (TCL) which is a structured programming language. Cache technology is used to mirror either CHARMM or CHARMM functions into TCL. The free format data structures and input/output abilities of GRACE allow it to interface with QM packages like CADPAC, GAMESS-UK or GAUSSIAN with its own systems or CHARMM. The most sophisticated built-in user interface is between GRACE and CHARMM. When GRACE is using CHARMM, CHARMM is run as a 'child process' to TCL. By routing all commands that are unrecognised by TCL, and all commands that start with a capital letter into CHARMM, the programming language of CHARMM is completely integrated into the TCL language.

CHARMM data structures, such as the position and gradient arrays are 'reflected' into TCL arrays. When these arrays are accessed in TCL, GRACE keeps track of whether the TCL arrays need updating from CHARMM. In the reverse direction, before CHARMM performs any task, GRACE will check to see whether it needs to update CHARMM's data structures from TCL. This means that the CHARMM data structures can be interfaced with the optimisers and analysis tools in GRACE. In the work where the AM1 QM/MM method in CHARMM has been used, GRACE has treated both the QM and MM atoms identically. It has been the job of CHARMM to generate the appropriate QM/MM function and gradient. The interfaces with CADPAC and GAMESS-UK is less sophisticated. It operates as part of the *ab initio* QM/MM modelling method in GRACE. When the energy and gradient of a QM/MM system is required, GRACE constructs an appropriate input deck for either GAMESS-UK or CADPAC, runs that code and parses out the energy and gradients from either the punch or the log file. The MM components of the QM/MM function are then constructed via the interface with CHARMM. None of this requires any interface with the QM/MM methodologies in CHARMM.

la - a stand alone interpreter

One of the programs launched by GRACE is the linear algebra algorithm or *la*. This code is a stand alone interpreter written in FORTRAN. It contains most of the linear algebra and optimisation methods plus some other functions. The TCL part of GRACE controls *la* by launching it as a 'child process' as it does CHARMM. TCL sends instructions to *la* as a simple alphanumeric data stream. *la* then returns TCL executable code which instructs the TCL interpreter to fill its arrays with information or to perform various functions, like acquiring energies and gradients. The transfer of many of GRACE's vital algorithms to *la* makes these functions available to programs written in other languages. This improves the long term viability of these algorithms in the event that TCL becomes obsolete.

in_core - information flow management

To avoid problems with compiling code written in C with that written in FORTRAN, the communication between CHARMM and TCL is handled using UNIX pipes. Under UNIX, FORTRAN unit 5 is always allocated to the standard input pipe and unit 6 to standard output. This method creates a problem of flow control. If TCL is trying to write to CHARMM's input whilst CHARMM is trying to write to its output, CHARMM's output pipe will not be read by TCL. This will cause the CHARMM process to be halted at which point it will perform no more reading, and so the TCL process will also be halted. *in_core* prevents this by acting as a memory buffer between TCL and CHARMM. TCL communicates to *in_core*'s standard input whilst in_core's standard output talks to CHARMM's standard input. *in_core* stores the information from TCL in memory until it is sent an instruction to flush this information to CHARMM. This allows TCL and CHARMM to run asynchronously. As *in_core* is a separate process to CHARMM, it could be written in C without compiler portability problems. C allows it to take advantage of dynamic memory management to store considerable amounts of data. Memory is allocated exponentially to avoid excessive calls to *malloc*.

Available optimisers

GRACE contains a wide range of optimisation algorithms. The TCL user interface makes them very flexible and easily interchangeable.

***lbfgs*:** A low-memory second-order optimiser using the same technology that is behind the 'work horse' ABNR optimiser in CHARMM.

***Powell*:** An auto-restarting conjugate gradients method implemented in TCL.

***P-RFO*:** Mode following optimiser based on the implementation by J. Baker. This extremely robust optimiser can search for minima, or other saddle points, whilst interfacing with GRACE's linear algebra systems to allow complex optimisation strategies.

As new optimisation algorithms and updates become available, they can be ported into GRACE.

Visualisation

Visualisation codes are not in-built components of GRACE; however, it can interface with programs such as XMol or RasMol via a series of pdb files that can be read and animated. This makes the visualisation of eigenvectors during saddle point location possible, ensuring that the correct mode is followed. The interface with RasMol is of a similar type to that with CHARMM. This allows a 'programmable' visualisation method. The interface lets the RasMol images be updated as the GRACE code runs which lets the evolution of the chemical model be viewed interactively. The interface with XMol is via the writing out of pdb files which XMol can then read. The major purpose behind this is to take advantage of the animation and measuring facilities in XMol. GRACE can produce a pdb file containing several structures which XMol can then show as an animation. This method has been used heavily for the viewing of eigenvectors of the hessian matrix during saddle point location.

An example GRACE script for a transition-state search

The script below is for the location of the transition-state for the bimolecular solvolysis reaction of ^{iso}-propyl chloride, $(CH_3)_2CHCl + H_2O \rightarrow (CH_3)_2CHOH_2^+ + Cl^-$.

```
source $env(GRACE_LIB)/grace.tcl                    Set location of GRACE files
set GLOBAL(scratch) ~/propyl                        Set scratch disk location
init i                                              Initialise GRACE
PRNL 6                                              Set CHARMM print and warning levels
WRNL 6
!-------- BEGIN RTF ---------                       Introduction of your CHARMM commands
Read rtf card                                       Note that at the start of a line, uppercase refers to CHARMM commands
* a simple rtf for propylcl + OH2 + TIP3P           and lowercase refer to GRACE commands
*
24  1                                   \
MASS   4 HT    1.000  ! H-bonding hydrogen          Information on atom types used in simulation
MASS  12 CT   12.010  ! Carbon
MASS  28 CL   35.000  ! Cl
MASS  75 OT   15.999  ! TIP3P oxygen

! TIP3P                                             TIP3P water information
RESI TIP3   0.000
GROUP
ATOM OH2 OT   -0.83400
ATOM H1  HT    0.41700
ATOM H2  HT    0.41700

BOND OH2 H1 OH2 H2 H1 H2
ANGLE H1 OH2 H2
ACCEPTOR OH2
PATCHING FIRST NONE LAST NONE

RESI PRO 0.0                                        Propyl Chloride +  nucleophilic water information such as:
GROUP                                               Name of residue and total charge on quantum system
ATOM  CR1  CT  -0.1  !      HM1  HM2  HM3          Atom labelling, atom type, specific atom charge and connectivity data for
ATOM  XCL  CL  -0.2  !           \  |  /            quick reference
ATOM  CM1  CT  -0.2  !  HO1      CM1
ATOM  CM2  CT  -0.2  !     \       |
ATOM  HD1  HT   0.1  !   OT1---- CR1 ---XCL
ATOM  HM1  HT   0.1  !     /       |  \
ATOM  HM2  HT   0.1  !  HO2    HD1  CM2
ATOM  HM3  HT   0.1  !               /  |  \
ATOM  HM4  HT   0.1  !           HM4 HM5 HM6
ATOM  HM5  HT   0.1  !
ATOM  HM6  HT   0.1  !
ATOM  OT1  OT  -0.5  !
ATOM  HO1  HT  0.25 !
ATOM  HO2  HT  0.25 !
END

Read para card
* Simple parameters propylcl + H2O
*

BONDS
HT  OT   450.0    0.9572
HT  HT     0.0    1.5139

ANGLES
HT  OT  HT    55.0   104.52

DIHEDRALS

NONBONDED nbxmod  5 atom cdiel shift vatom vdistance vswitch
CUTNB 40.0 ctofnb 38.0 ctonnb 36.0 eps 1.0 e14fac 1.0 wmin 1.0

 CT   0.000000  -0.080000   2.060000 ! All carbons
 HT   0.000000  -0.046000   0.224500 ! H bond H
 CL   0.000000  -0.126620   2.144580 ! Chlorine
 OT   0.000000  -0.152100   1.768200 ! TIP3P OXYGEN
END                                                 End of CHARMM structure and parameter data

Read sequence card                                  Read in sequence information
* simple
```

```
*
1
PRO
Generate PRT setup

Read sequence TIP3 494
Generate SOLV setup noangle nodihedral

Open unit 10 read form name propyl_ts_inp.pdb
Read coor pdb unit 10
Close unit 10

Nbonds atoms switch cutnb 40.00 ctofnb 39.00 ctonnb 35.00
Updat inbfrq 1000 ihbfrq 0 cutnb 999.0 ctonnb 997.0 ctofnb 998.0 Cdie

Quan sele resi PRO end am1 charge 0

Prnl 0
Wrnl 0

wait


proc frelax xvec\ gvec {



upvar $xvec xparam $gvec grad
global X Y Z WMAI DX DY DZ PARA QMSTART

qmmm silent unmake_pos=xparam
  Scal xref stor 1
  Scal yref stor 2
  Scal zref stor 3
  Cons fix  sele  core end
  Scal zref reca 3
  Scal yref reca 2
  Scal xref reca 1
  Prnl 5
  Mini abnr nstep 5000 tolg 0.005 step 0.0001 nprint 1000
  Prnl 0
  Scal xref stor 1
  Scal yref stor 2
  Scal zref stor 3
  Cons fix  sele  none end
  Scal zref reca 3
  Scal yref reca 2
  Scal xref reca 1
  set X(1)

set f [fcharmm xparam grad]


Quick 1 12
set d1 $PARA(DIST)
Quick 1 2
set d2 $PARA(DIST)
puts "+++ Distance 1-12 = $d1"
puts "+++ Distance 1-2  = $d2"
flush stdout
return $f }

qmmm \
  fmm=all


Prnl 5
Defi core sele resi PRO end
Prnl 0
acquire_selection selection=core var=lcore
puts "Selected atoms and types"
foreach i $lcore {
puts "$IUPC($i) $i $RESI($i) $RESN($i)"
}
flush stdout
```

The system contains only two segments PRT (QM) and SOLV (MM) made up from the PRO and TIP3P residues, respectively

Now read in the sequence for the 494 TIP3P water molecules

Now read in the co-ordinates of the MM & QM atoms from an external pdb file

Set non-bonding interactions and proximity warnings

Select residue that will be the QM system, set the charge (0) and computaional method (AM1)

Wait for CHARMM to do it bit before we move on, or the output of grace and CHARMM get confused

Set up procedure to perform relaxed environment functions. This procedure computes the f and g from fcharmm and then minimises the rest of the system around it, you must make the defined 'core' selection to represent the part of the system that is in the hessian, for this procedure to work

Update the position and gradient vectors to xparam and grad

Always relax environment totally first set up core positions use the silent option so it does not print out lots of useless information. Start the initial environment optimisation using the ABNR optimiser.

Get energy and gradient, the energy will be the total energy, but the gradient vector will only span the sub set defined by the QM/MM statement

Ask GRACE to instruct CHARMM to print out the bond lengths between bonds of interest

Define the parts of the QM/MM system. Grace does not 'know' if you have parts of the CHARMM system AM1 or not, so first make everything fixed MM

Here we make the 'core' definition and set Prnl high so we can see if it has worked

```
  Scal xref stor 1
  Scal yref stor 2
  Scal zref stor 3
  Cons fix  sele  core end
  Scal zref reca 3
  Scal yref reca 2
  Scal xref reca 1
  Prnl 5
  Mini abnr nstep 5000 tolg 0.001 step 0.0001 nprint 20
  Prnl 0
  Scal xref stor 1
  Scal yref stor 2
  Scal zref stor 3
  Cons fix  sele  none end
  Scal zref reca 3
  Scal yref reca 2
  Scal xref reca 1
  set X(1)

qmmm \
  mm=core \
  make_pos=pos \
  start
gethess \
  func=fcharmm \
  position=pos \
  dump_name=$GLOBAL(scratch)/hess.dump \
  hessian=h
la symm matrix=h result=h
la eign matrix=h
pmatrix matrix=values

optimise \
  position=pos \
  func=frelax \
  nfunc=100000 \
  omin=0.800 \
  osmin=0.001 \
  rmin=0.90 \
  rmax=1.10 \
  step=0.001 \
  maxstep=0.01 \
  mode=1 \
  lock \
  dump_name=$GLOBAL(scratch)/ts.dump \
  method=ef \
  hessian=h \
  toler=0.3

xyz>p x=X y=Y z=Z p=final_pos
array>list array=final_pos list=lp
array>list array=h list=lh
la eign matrix=h
array>list array=roots list=lr

Open unit 67 write form name propyl_ts.pdb
Write coor pdb unit 67
* pdb of the transition-state for propylcl + H2O
*
Close unit 67

exit
```

Now make the core bit movable MM and then generate a position vector for it, so GRACE 'internally' knows the parts that are movable.

Carry out an energy minimisations using the ABNR optimiser

Diagonalise martix

Print out roots values

Perform an EF search for a transition-state

Follow mode 1, hopefully this is the right mode other wise we can stop the simulation, change the mode to the correct one and then restart it.
'dump' can be used to restart the simulation or used to view the mode that is being followed.

Write out a pdb file of the final transition-state structure

All done.